

**UNIVERSIDAD NACIONAL AUTÓNOMA DE
MÉXICO**

**ESCUELA NACIONAL DE ESTUDIOS PROFESIONALES
ARAGÓN**

**DESARROLLO E IMPLEMENTACIÓN DE UN
DIGITALIZADOR DE VOZ UTILIZANDO EL
MICROPROCESADOR TMS3477**

TESIS QUE PARA OBTENER EL GRADO DE:

INGENIERO MECÁNICO ELECTRICISTA

PRESENTA:

QUINTANA ROCHA, HUGO

**TESIS CON
VALIA DE ORIGEN**

**TESIS CON
FALLA DE ORIGEN**

1996



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.



UNIVERSIDAD NACIONAL
AUTÓNOMA DE
MÉXICO

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
CAMPUS ARAGÓN

UNIDAD ACADÉMICA

53
24

Ing. RAÚL BARRÓN VERA
Jefe de la Carrera de Ingeniería
Mecánica Eléctrica,
Presente.

En atención a la solicitud de fecha 9 de agosto del año en curso, por la que se comunica que el alumno HUGO QUINTANA ROCHA, de la carrera de Ingeniero Mecánico Electricista, ha concluido su trabajo de investigación intitulado "DESARROLLO E IMPLEMENTACIÓN DE UN DIGITALIZADOR DE VOZ, UTILIZANDO EL MICROPROCESADOR TMS3477", y como el mismo ha sido revisado y aprobado por usted se autoriza su impresión; así como la iniciación de los trámites correspondientes para la celebración del examen profesional.

Sin otro particular, le reitero las seguridades de mi distinguida consideración.

ATENTAMENTE
"POR MI RAZA HABLARA EL ESPÍRITU"
San Juan de Aragón, Mex., 14 de agosto de 1996.
EL JEFE DE LA UNIDAD


LIC. ALBERTO IBARRA ROSAS

~~ccp~~ Asesor de Tesis.
~~ccp~~ Interesado.

AIR/lla.

DEDICATORIA

A MI ESPOSA E HIJA.

FRANCISCA ROBLES

FANNY QUINTANA

Gracias, por soportar privaciones
y desconsuelos, por mantener su confian-
za y su cariño en mi, por continuar jun-
to a mi por todo lo que hemos conseguido.

AGRADECIMIENTO

UN AGRADECIMIENTO MUY ESPECIAL PARA EL LOGRO DE ESTA
TESIS AL ING: SILVIA VEGA.

GRACIAS.

INDICE

	PÁGINA
PROLOGO	
INTRODUCCIÓN	i
CAPITULO 1	
GENERALIDADES	1
1.1. Ventajas de la comunicación digital ante la comunicación analógica.	2
1.2. Técnicas de modulación por pulsos.	4
1.3. Sistemas de comunicación PCM generalizado.	6
1.4. Técnicas digitales para implementar síntesis de voz.	11
1.4.1. Tamaño de las unidades de voz.	12
1.4.2. Propiedades básicas de voz.	13
1.5. Métodos y Técnicas de síntesis de voz.	16
1.5.1. Codificación de forma de onda.	16
1.5.2. Codificación de la fuente (Vocoding) o Paramétricas.	18
1.5.3. Codificación híbrida.	21
1.5.4. Sistemas de texto ilimitado (Texto a Voz) y Sistemas de texto limitado (Respuesta Sonora).	22
1.6. Calidad de la Voz.	23
1.7. Estado actual de la tecnología para implementar síntesis de voz.	24

1.8. Compromiso entre complejidad y funcionamiento.	28
---	----

CAPITULO 2

FUNDAMENTOS TEÓRICOS	30
2.1. El teorema de muestreo de Nyquist.	32
2.1.1. Muestreo ideal.	32
2.1.2. Muestreo real.	35
2.1.3. Reconstrucción de la señal analógica.	37
2.2. Requerimientos de codificación.	40
2.2.1. Niveles de cuantización.	40
2.2.2. Ancho de banda.	42
2.3. Ruido de cuantización.	44
2.3.1. Relación señal a ruido de cuantización en niveles igualmente espaciados.	44
2.3.2. Relación señal a ruido en señales con compansión.	48
2.4. El sistema Bell System T1 carrier.	53
2.4.1. Funcionamiento.	53
2.4.2. Repetidores regenerativos.	55
2.5. Proceso de compresión y expansión (companding) de una señal.	56
2.5.1. Cuantización logarftmica o leyes de compandig.	57
2.6. Efecto del ruido en el canal de transmisión.	66
2.6.1. Características del ruido.	66
2.6.2. Probabilidad de error.	68

2.6.3. Capacidad de canal.	72
2.7. Filtros Pasa - Bajas.	76
2.7.1. Filtro activo PB - VCVS.	76

CAPITULO 3

TÉCNICAS DE MODULACIÓN DIGITAL	80
3.1. Técnicas de compresión de señales de forma de onda.	81
3.2. Muestreo y cuantización de señales de voz.	82
3.3. Modulación por codificación de pulsos (PCM).	85
3.3.1. PCM lineal.	85
3.3.2. PCM logarítmico.	88
3.3.3. Modulación Delta (DM).	90
3.3.4. Modulación Delta Adaptiva (ADM).	96
3.4. Comparación de los sistemas PCM logarítmico, DM y ADM.	99
3.5. Modulación Delta con Pendiente que Varía Continuamente (CVSD).	101
3.6. Codificación en sub-bandas (SBC).	102
3.7. Técnicas de modulación en amplitud.	106
3.8. Técnicas de modulación en frecuencia.	107
3.9. Técnicas de modulación en fase.	110
3.10. Técnicas de modulación híbridas.	114
3.11. Elementos de juicio para seleccionar una técnica de modulación.	115
3.12. El sistema transmisor - receptor.	116

CAPITULO 4

Reporte de Aplicación del TMS3477	117
Consideraciones.	118
Introducción.	118
4.1. Funcionalidad del circuito.	121
4.1.1. Estación de GRABACIÓN / REPRODUCCIÓN.	121
4.1.2. Estación de REPRODUCCIÓN.	122
4.2. Operación del circuito. DIGITAL	122
4.2.1. Conversión de Direcciones.	123
4.2.2. Multiplexaje de Datos.	123
4.2.3. Memoria de Escritura / Lectura.	125
4.2.4. Organización de Resistencias.	126
4.3. Operación del circuito. ANALÓGICA	
4.3.1. Línea de Entrada.	128
4.3.2. Línea de Salida.	128
4.3.3. Bocina.	130
4.3.4. Micrófono.	130
4.3.5. Oscilador.	132
Conclusiones.	134
Bibliografía.	135
Apendice.	

PROLOGO

El presente trabajo además de tener el objetivo de obtener el título de Ingeniero Mecánico Electricista, es una introducción a los sistemas de procesamiento digital de señales.

El trabajo sigue una secuencia para poder entender más fácilmente lo que significa el procesamiento de señales y su aplicación. Además el desarrollo de las comunicaciones digitales en los últimos 30 años, ha investigado varios aspectos interesantes en lo que respecta a costo, facilidad de implementación y funcionalidad.

Comenzamos con la historia de la digitalización de señales a través de un sistema generalizado de PCM (Modulador por Código de Pulsos) y nos adentramos a los requerimientos fundamentales para implementar síntesis de voz, así como sus métodos y técnicas. Concluimos con las expectativas a futuro del uso de digitalizadores de voz aplicados a funciones de la vida diaria (Capítulo 1).

Teniendo conocimiento de las técnicas y síntesis en el desarrollo de sistemas digitales. Nos corresponde ahora entender que para todo sistema se requiere contar con las bases teóricas que permitan establecer las especificaciones del funcionamiento, comparar con otros dispositivos similares, o en el mejor de los casos desarrollar nuevas fórmulas para la teoría en estudio (Capítulo 2).

La separación física entre emisor y receptor implica la existencia de un medio de transmisión por medio del cual viaje la información. Este medio o canal de transmisión puede ser un par de hilos, cable coaxial, guía de onda, el espacio abierto, etc. En los cuales es necesario aprovechar eficazmente: sus componentes, la técnica utilizada y en el mejor de los casos una opción a un bajo costo. Conocer las más relevantes será lo que se discute en este capítulo (Capítulo 3).

Actualmente una técnica de compresión de señales desarrollándose ampliamente es el CVSD (Deltamodulación con Pendiente que Varía Continuamente), por lo que su utilización viene siendo una realidad. Así que teniendo la facilidad de tener esta técnica en un chip puede hacerse realidad un digitalizador de voz a bajo costo. Además de ser factible la realización de una interface para el almacenamiento de datos (Capítulo 4).

INTRODUCCION

Hoy en día el avance tecnológico ha hecho que las computadoras en la sociedad moderna sean ampliamente utilizadas. El procesamiento digital de voz promete revolucionar los caminos de interacción del hombre con las máquinas, es decir, dotando a las máquinas con la habilidad para interpretar y actuar por medio de instrucciones habladas. Dado ésto, el reconocimiento de voz, se está convirtiendo en una herramienta indispensable para poder lograr la automatización. La síntesis de voz, que dá voz a un dato digital, está progresando rápidamente con lo que se liberará a los usuarios del uso de las pantallas y teclados. Por tanto, el desarrollo de las técnicas para procesar la voz han ido en aumento.

EL HABLA COMO MEDIO DE COMUNICACION.

Los seres humanos utilizan principalmente el habla como medio de comunicación , siendo éste el único ser viviente que ha adquirido y usado el habla con éste fin, y para producir la voz ha utilizado el aparato fisiológico que tiene como propósito vital, comer y respirar. Esto ha sido posible debido a que se encuentra envuelto en una atmósfera de aire, con lo que se han tenido las condiciones para aprender a emitir señales producidas por vibraciones longitudinales en el aire, por tanto, la voz consiste en rápidas y determinadas fluctuaciones en la presión del aire. El proceso de comunicación por medio del habla se lleva a cabo cuando los sonidos y presiones del aire son generados y emitidos por el aparato vocal del orador y captados por el oído y transmitido al cerebro para su comprensión por el escucha fig.1.

Las ondas acústicas producidas por el hombre son las vibraciones longitudinales producidas cuando la cavidad torácica se expande y contrae para expulsar el aire desde los pulmones hacia afuera a través de la traquea al glotis. Si las cuerdas vocales están tensadas, el flujo de aire producirá vibración en las cuerdas produciendo pulsos cuasi-perfódicos. Si las cuerdas vocales son extendidas separadamente, el aire pasa fluyendo a través del glotis produciendo un flujo de aire turbulento cuasi-aleatorio. El aire fluyendo a través de la faringe pasa al tracto vocal, dependiendo de la posición de los articuladores (labios, lengua, etc.) y del escotillón del paladar (paladar suave), el aire fluirá por la boca, por la nariz o por ambas y la voz será emitida.

- 1o.- Formulación del Mensaje
- 2o.- Código del Lenguaje
- 3o.- Acción Neuro-Muscular
- 4o.- Cuerdas Vocales
- 5o.- Tracto Vocal

- 1o.- Oído
- 2o.- Transductor Neural
- 3o.- Código del Lenguaje
- 4o.- Comprensión del Mensaje

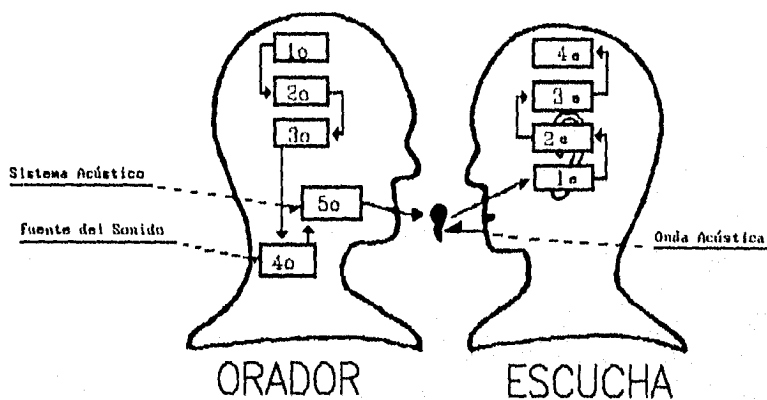


Fig.1 Proceso de comunicación por medio del habla.

En la fig.2 y 3 se muestra un diagrama esquemático de la producción de voz humana y el plano sagital de rayos X de la cabeza humana, donde es posible apreciar el tracto vocal que es un tubo acústico no uniforme que se extiende desde la glotis a los labios, con una longitud de alrededor de 17 cm. en un adulto. Este es de una área de sección transversal no uniforme que depende de la posición de los articuladores y varía desde cero (completamente cerrada) a cerca de 20 cm. El tracto nasal se extiende desde el paladar hasta las narices, teniendo una longitud aproximada de 12 cm. y un volumen de 60 cm.

El progreso y desarrollo del habla humana desde su forma primitiva con gruñidos y chillidos hasta el estado donde se encuentra con código acústico especializado ha sido a través de un período largo de tiempo, pues se especula que los primeros hombres se comunicaban por medio de las señales de las manos, con lo que una de las hipótesis ampliamente difundidas es de Sir Richard Paget " El hombre manejó la invención del habla cuando sintió la necesidad de expresar sus pensamientos teniendo la dificultad de hablar con sus manos." A través del paso de la historia, la producción del habla artificial ha sido de gran interés para el hombre, por lo que han surgido diversas formas de generarla, como puede apreciarse en las antiguas civilizaciones griega y romana en donde existieron estatuas que hablaban, voces milagrosas y oráculos en donde la voz era emitida por el encargado del templo o por el sacerdote y transmitida por medio de un tubo fonético oculto. Con el paso del tiempo, los cambios sociológicos y culturales dieron paso a mecanismos con fuelle. El intento de imitar voz humana no ha cesado y actualmente existen una gran cantidad de técnicas para poder realizarlo.

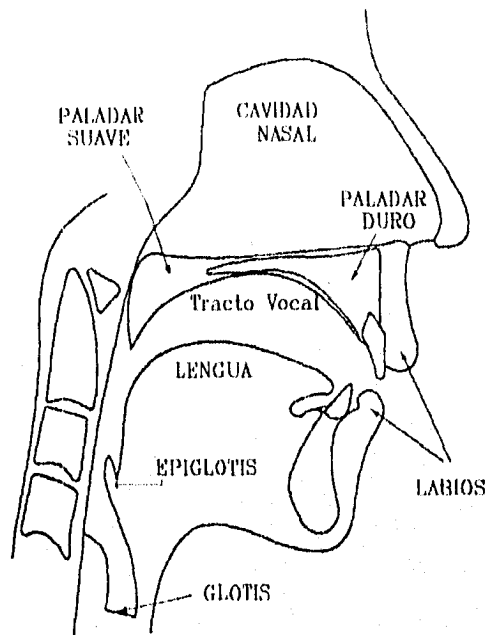


Fig.2 Plano sagital de rayos X de la cabeza humana.

La comunicación es el proceso por medio del cual la información se transfiere de la fuente al destino, por tanto existen sistemas de comunicación que son un conjunto de mecanismos que proporcionan el enlace para la información entre destino y fuente (transmisor, canal de transmisión, receptor). Para caracterizar la voz en la comunicación se pueden mencionar dos formas; la primera, es una idea introducida por Shannon, en donde la voz es representada en términos del contenido de un mensaje o información (por ejemplo, 10 fonemas por segundo) llamada **TEORIA DE LA INFORMACION**.

La segunda forma es caracterizando la voz en términos de **SEÑAL** (onda acústica producida por voz o música) llevando la información del mensaje. La información contenida en la voz es intrínsecamente de naturaleza discreta, es decir, está compuesta de una secuencia de sonidos, los cuales pueden ser clasificados fonemas los que son una concatenación de elementos. Estos sonidos y la transición entre ellos sirve como una representación de la información. El arreglo de dichos sonidos son gobernados por las reglas del lenguaje. El estudio de dichas reglas y sus implicaciones en la comunicación humana es del dominio de la lingüística, y el estudio y la clasificación de los sonidos de la voz se llama fonética. El aspecto técnico del sistema de comunicaciones trae como consecuencia una variedad de representaciones de la señal de voz, sin embargo, existen dos características importantes que se deben mantener en cualquier sistema :

1.- Conservación del contenido del mensaje.

2.- La representación de la señal de voz debe ser una forma conveniente para su almacenaje o transmisión.

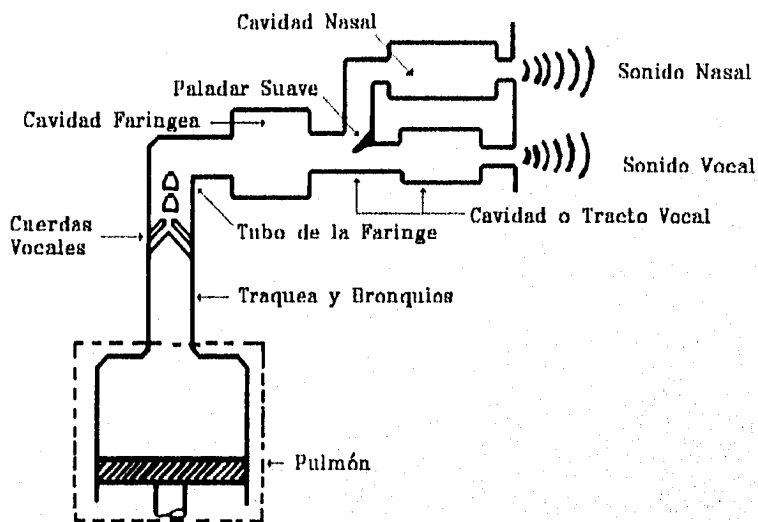


Fig.3 Esquema del mecanismo de producción de voz.

PROCESAMIENTO DIGITAL DE SEÑALES.

El problema general del procesamiento y manipulación de la información involucra dos aspectos importantes: la representación de la señal y la transformación de la señal en una forma conveniente. Por ejemplo, un sistema de identificación de una persona en forma automática de un conjunto de oradores, podría usarse una representación espectral en función del tiempo de la señal de voz. Una transformación posible de la señal puede ser el promedio espectral a través de una palabra o frase, la cual se comparará con el almacenado, y entonces basada sobre la medida espectral similar elige la identidad del interlocutor. Para la señal de voz, la representación se puede encontrar en cualquiera de sus tres formas :

Forma de onda.

Forma paramétrica o de la fuente.

Forma híbrida.

Las técnicas digitales son ampliamente utilizadas para el procesamiento de señales dado que el desarrollo tecnológico, el costo y la flexibilidad de los sistemas y computadoras digitales sobrepasa a los sistemas analógicos. Para el procesamiento de voz existen muchas razones para la utilización de técnicas digitales, entre ellas, la voz en forma digital puede ser transmitida en un canal ruidoso sin degradaciones, o ser transmitida simultáneamente con datos en una malla sin necesidad de distinción entre ellas, exceptuando en la decodificación.

PROCESAMIENTO DIGITAL DE LA VOZ.

El desarrollo del área del procesamiento digital de señales se ha llevado a la realidad debido a la realización del procesamiento digital de la voz. Los problemas generales en donde se ve involucrado el procesamiento digital de la voz puede dividirse en tres clases:

Análisis de Voz.
Síntesis de Voz.
Análisis - Síntesis de Voz.

Cuando nos interesamos únicamente en el ANALISIS DE VOZ, que es simplemente el proceso de estimación de los parámetros (variables en el tiempo) del modelo para producir voz, desde la señal de voz que es asumida para ser la salida del modelo. Los sistemas de análisis de voz sirven usualmente como la "entrada final" para sistemas automáticos de reconocimiento de voz y para sistemas de verificación o identificación automática de un orador. Cuando nos interesa la SINTESIS DE VOZ, que es el proceso de producción de una señal acústica, controlando el modelo de producción de voz con un conjunto apropiado de parámetros. Los sistemas de síntesis de voz son usualmente la "salida final" en forma de voz de una computadora o de una máquina lectora de texto escrito. La tercera clase involucra primero el análisis seguido de la síntesis de voz que es la más utilizada en las aplicaciones del proceso digital de la voz. La calidad de los datos extraídos en la forma hablada dependerá de las técnicas utilizadas en el procesamiento, así como de la sofisticación de los codificadores y decodificadores, por tanto, si el modelo es lo suficientemente preciso y los parámetros son perfectamente determinados, la salida sonora será indistinguible de la voz natural.

Cuando se tiene determinada la necesidad y el uso del sistema es cuando sabremos en cual de los tres problemas se está involucrado, así como que método de procesamiento digital de señales será utilizado para representar, procesar y almacenar o transmitir la voz. En el siguiente cuadro se muestran algunas aplicaciones típicas de la utilización del procesamiento digital de la voz:

Aplicaciones del procesamiento Digital de la Voz.

Codificación de Voz.
Síntesis de Voz.
Verificación e Identificación de un orador.
Reconocimiento de Voz.
Ayuda a Impedidos.
Transmisión de Datos.

CODIFICACION DE VOZ.

Las técnicas de codificación de voz están siendo ampliamente estudiadas y se han ido realizando mejoras para lograr un uso más eficiente del ancho de banda del canal de transmisión. También se desarrollan para lograr el almacenaje de más información en menos memoria.

SINTESIS DE VOZ.

Los sistemas de síntesis de voz han sido ampliamente estimulados y han logrado rápidos avances debido a la necesidad de un almacenamiento de voz para las computadoras con sistemas de respuesta sonora, para la reproducción de un mensaje previamente codificado y almacenado que puede ser extraído por una persona desde una terminal o teclado, o de un sistema telefónico, o bien de un juguete. Por ser un proceso digital tiene las siguientes ventajas sobre las técnicas analógicas:

- a) La reproducción de un mensaje es de una calidad equivalente al original.
- b) No existe desgaste mecánico de los componentes.
- c) El mensaje se reproduce siempre desde su inicio.
- d) Bajo costo.

VERIFICACION E IDENTIFICACION DE UN ORADOR.

La verificación e identificación involucra obtener la autenticidad o identificación de un orador de un grupo grande de posibles oradores, así como en los sistemas de verificación se debe decidir si el orador es la persona que demanda ser. Esto es potencialmente aplicable en requerimientos de control de acceso de información y en áreas restringidas, al igual que en varias clases de transacciones crediticias automáticas.

RECONOCIMIENTO DE VOZ.

El reconocimiento de voz es una área multidisciplinaria que es fuertemente dependiente de las restricciones establecidas sobre el orador. Este tipo de sistemas son muchos y muy variados, tales como computadoras que responden a la comunicación sonora y máquinas de escribir operadas por voz.

AYUDA A IMPEDIDOS.

Una variedad de técnicas de procesamiento de señales están siendo aplicadas para aplicaciones didácticas para facilitar el aprendizaje en general, y para dar facilidades a personas impedidas como ciegos y minusválidos.

TRANSMISION DE DATOS.

Debido al crecimiento y demanda para la transmisión de datos en general y principalmente utilizando la red telefónica y al desarrollo de la comunicación vía satélite, micro-ondas y sistemas de comunicación óptico, ha sido necesario que surgan sistemas que digitalizan la voz a bajas razones de bits tanto como sea posible.

CAPITULO 1

GENERALIDADES.

1.1.VENTAJAS DE LA COMUNICACION DIGITAL ANTE LA ANALOGICA.

El desarrollo de las comunicaciones digitales en los últimos 30 años, ha permitido superar en grado sumo al status de las comunicaciones analógicas, que de una u otra forma han permanecido estancadas da la moda en la investigación por las técnicas digitales. Varios aspectos interesantes en lo que respecta a costo, facilidad de implementación y funcionalidad, nos conducen a una comparación entre ambos tipos de transmisión.

A) COSTO

El gran impulso que tuvieron los circuitos integrados desde su inicio, logró que el costo de estos dispositivos haya descendido a un nivel bastante adquisitivo. Los sistemas digitales basados en esta tecnología por consiguiente resultan más económicos, no así los sistemas analógicos que requieren material adicional a los semiconductores, a los que podríamos llamar tradicionales, para solventar los limitantes de este tipo de transmisión, como serían relevadores, cable pupinizado, etc. De hecho en lo que a inversiones se refiere, la industria de las telecomunicaciones tiende a elegir la digital. En la fig.1.1. se muestran los pronósticos de inversión para distintas ramas de las comunicaciones en E.U. para 1990.

B) FACILIDAD DE IMPLEMENTACION

Toda vez que la circuiteria digital permite ser presentada con módulos independientes e interconectables entre sí controlados por un procesador central, ha sido factible instalar, expandir o trasladar todo un sistema PCM con suma rapidez. Esto ha traído la tendencia en la telefonía al desarrollo de conmutadores digitales modulares para n canales variables. Por lo regular los sistemas PCM necesitan conectarse punto a punto o vía repetidores regenerativos. Hasta el momento las conexiones entre centrales terminales analógicas requieren una gran cantidad de cables, tanto para control y supervisión como para información en sí, lo cual trae consigo lentitud en la instalación de este medio de transmisión. Con el sistema PCM la conexión entre puntos terminales requiere de un solo canal de transmisión al través de un sistema multiplexado, por lo que regularmente se utilizan medios de transmisión ya instalados pensados originalmente para señales analógicas.

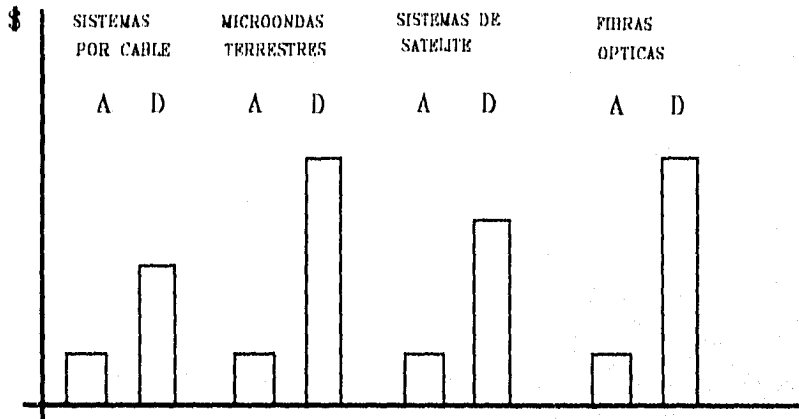


Fig.1.1. Pronosticos de Inversión para 1990 en E.U.

C) FUNCIONALIDAD

Debido a que en los sistemas digitales se trabaja con dos estados lógicos 1 ó 0, se puede tolerar un alto nivel de ruido antes de que se produzcan errores en la recepción. El valor de un dato físico expresado como palabra digital binaria, se determina con la presencia o ausencia de pulso en cada uno de los dígitos de la palabra. En el caso de señales analógicas, por ser del tipo continuo, el ruido a cualquier nivel de intensidad, distorsiona la información en proporción a la relación S/N. Si la línea de transmisión esta provista de repetidores a lo largo de la ruta, las señales digitales, se regeneran sin ningún ruido en cada punto repetidor intermedio; en contraste, el ruido de las señales analógicas se va acumulando al pasar de uno a otro repetidor hasta el punto de recepción final.

Los sistemas digitales presentan gran versatilidad ya que manejan información de varias naturalezas sin distinción entre ellas una vez codificadas, como sería voz, facsímil, video, etc., sin interferir unas con otras, lo cual permite agruparlas en serie, por paquetes regularmente, para su envío por un único medio de transmisión: un par de cables, cable coaxial, radio, fibra óptica, etc. Para lograr esto con los sistemas analógicos, se requieren equipos de multiplex en frecuencia, los cuales son caros y voluminosos, además de problemas por envejecimiento como son la inestabilidad de frecuencia en los osciladores de onda portadora, o corrimientos en la banda de paso diseñada originalmente para los filtros selectores.

1.2. TÉCNICAS DE MODULACION POR PULSOS

En el proceso de convertir una señal analógica en forma digital, el uso del muestreo en una señal durante cortos intervalos de tiempo ha desarrollado una serie de técnicas para modulación por pulsos. Las muestras indicativas del valor instantáneo de la información en el punto de muestreo, son representadas por uno o varios pulsos que están modulados de acuerdo con la amplitud de la información en esos instantes.

Existen cuatro métodos básicos de modulación por pulsos:

- 1. Modulación por amplitud de pulso (PAM).

Para este método la amplitud de un punto representa el valor instantáneo de la señal de entrada.

- 2. Modulación por ancho de pulso (PWM).

En este tipo de modulación, se varía la duración del pulso para representar un valor muestra.

- 3. Modulación por posición de pulso (PPM).

En este sistema cada pulso tiene una posición de referencia cero para una amplitud específica de la información. El valor instantáneo de la señal analógica en los puntos de muestreo aleja o acerca el pulso de su posición de referencia.

- 4. Modulación por codificación de pulsos (PCM).

En este sistema el valor muestreado de la señal analógica es codificado en un arreglo de dígitos binarios. Estos dígitos tienen la misma forma y amplitud, de esta manera el equipo receptor requiere únicamente detectar la presencia o ausencia de pulsos. Un sistema PCM es menos sensible al ruido y distorsión que las otras técnicas mencionadas fig. 1.2.1.

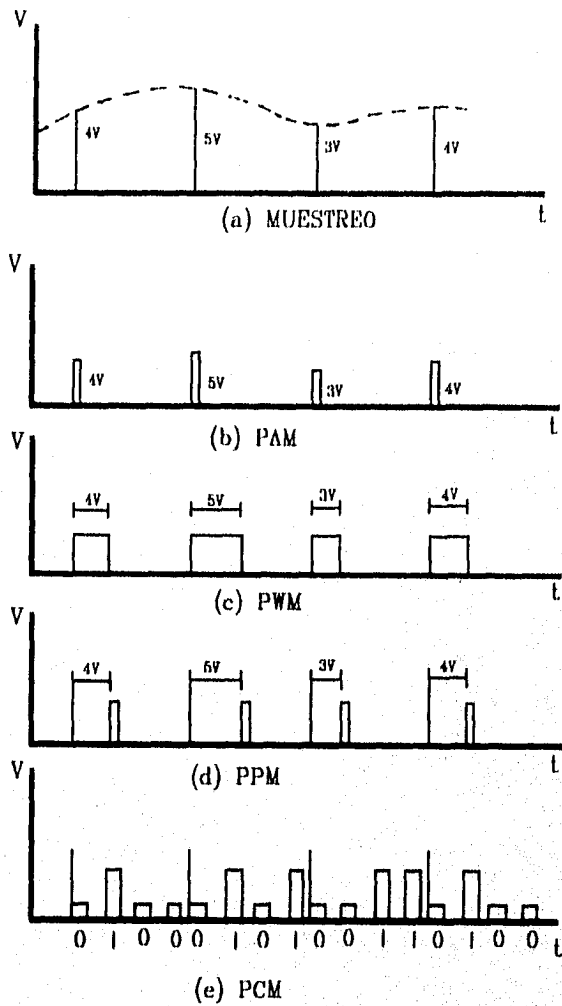


Fig.1.2.1. Tipos de Modulación por Pulsos.

1.3. SISTEMAS DE COMUNICACION DIGITAL PCM GENERALIZADO.

Aún cuando el equipo de comunicación para este sistema varía por aplicación específica, los bloques que se muestran en la fig.1.3.1. aparecen en la mayoría de los casos.

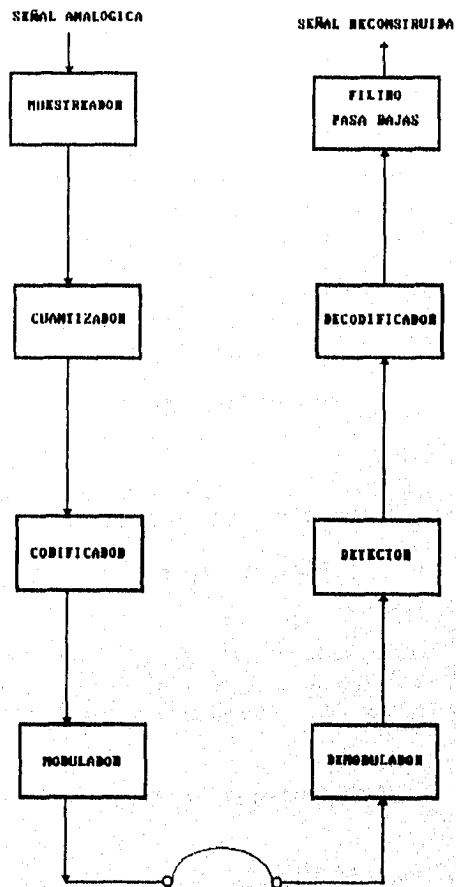


Fig.1.3.1. Sistema PCM Generalizado.

A) MUESTREADOR

El teorema de muestreo de Nyquist establece que una señal limitada en banda que no contiene componentes espectrales no mayores que la frecuencia F_m [Hz], está determinada en forma única por sus valores instantáneos en intervalos uniformes de $1/2F_m$ segundos. Es decir, que con muestras discretas de la señal analógica tomadas a intervalos regulares, tenemos la información suficiente para reconstruir la señal. Por lo tanto, el muestreador es un dispositivo que trabaja a una frecuencia mayor o igual a $2F_m$ [Hz] y obtiene el valor instantáneo de la señal en el instante que se realiza el muestreo.

B) CUANTIZADOR

En el argot digital, el redondeo de una magnitud al valor más próximo a un juego de valores discretos se denomina cuantización, esto es, en el instante de muestreo los pulso modulados en amplitud pueden tomar un cierto valor dependiendo de la amplitud de información; sería ideal recibirla de esta forma, desgraciadamente el ruido presente en todo canal de transmisión hace que el receptor tome una señal ligera o bastante distorsionada. Es por esto que la señal PAM se aproxima a un número finito de valores establecidos por el diseñador, de manera que si la separación entre niveles es mayor a la amplitud del ruido, el receptor tendrá la posibilidad de distinguir el nivel que ha sido transmitido. Ha resultado más eficiente un bloque antes de transmitir la información, el codificador.

C) CODIFICADOR

El codificador representa las muestras cuantizadas de la señal analógica por un código digital, como el binario, BCD, gray, etc. A la salida de este dispositivo tenemos la señal PCM.

D) MODULADOR

En el proceso de transmitir los pulsos codificados en un medio adecuado de transmisión, es necesario acoplarlos a las características de éste. Para ello, la transmisión se efectúa con ondas senoidales de mayor frecuencia que los pulsos portadores de información, que varían su amplitud, frecuencia o fase en correlación.

E) CANAL DE TRANSMISION

Es el medio físico por el que la información se traslada del sistema emisor al sistema receptor, y puede ser un par de cables, guía de onda, fibra óptica e inclusive el espacio libre.

F) DEMODULADOR

Este dispositivo extrae las variaciones en los parámetros de la onda senoidal que son representativas del valor de los dígitos binarios. Esto se logra por lo regular con un detector de envolvente para el caso de AM o por la multiplicación en el receptor con una frecuencia de valor igual a la de la portadora para el caso de FM.

G) DETECTOR

Elemento del sistema que compara la magnitud de cada pulso binario recibido con un nivel de decisión, e indica el valor del pulso decidido "1" si sobrepasa el umbral de referencia y "0" si no hay sobrepaso. A la salida se tiene una señal limpia, sin ruido, aunque tal vez con uno o varios bits erróneamente detectados debido a efectos indeseables en el canal de transmisión.

H) DECODIFICADOR

Realiza un proceso inverso a la codificación, es decir, proporciona a su salida una señal discreta cuantizada a un nivel de voltaje, dada la palabra digital presentada en su entrada.

I) FILTRO PASA BAJAS

Elimina las altas variaciones de los pulsos discretos de la señal decodificada, siguiendo la envolvente de las amplitudes de cada uno de ellos, obteniéndose con esto la señal analógica original reconstruida.

La fig.1.3.2. y fig.1.3.3. ilustra el proceso de la señal de información para su envío y recepción.

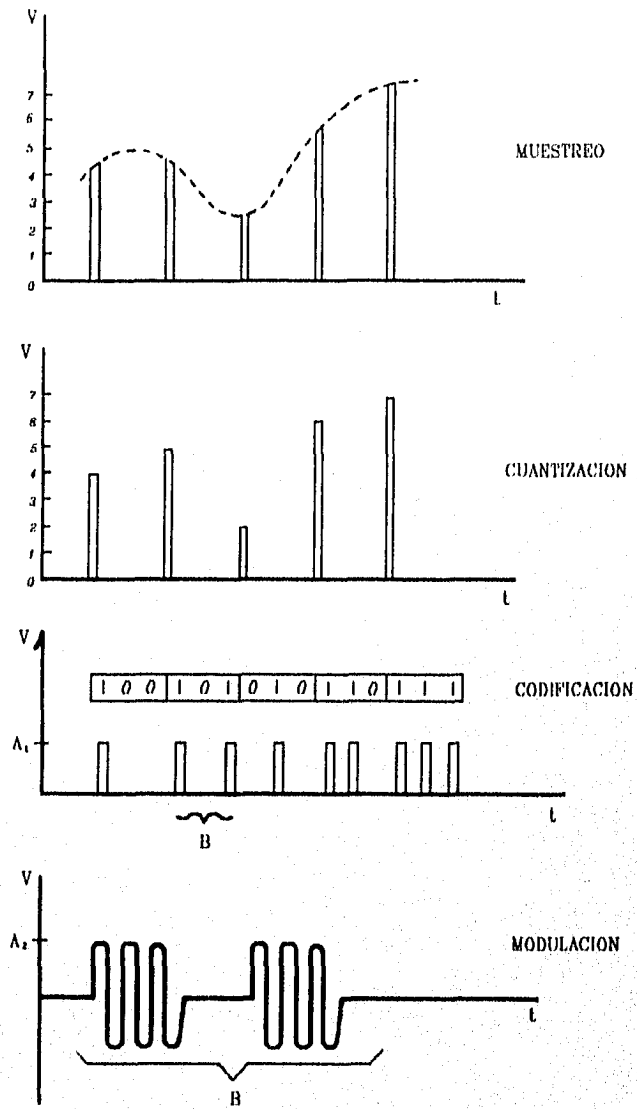


Fig.1.3.2. Proceso de Transmisión para un sistema PCM.

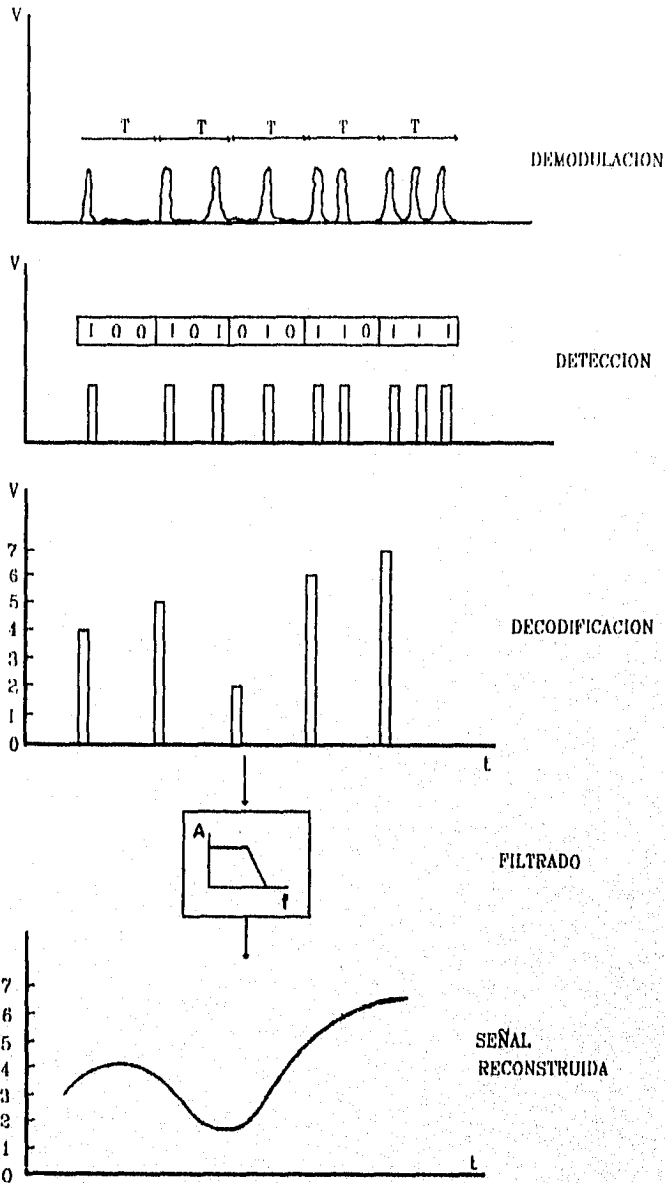


Fig.1.3.3. Proceso de Recepción para un sistema PCM.

1.4. TÉCNICAS DIGITALES PARA IMPLEMENTAR SINTESIS DE VOZ.

En años recientes, se ha desarrollado rápidamente la generación automática de formas de onda originadas por datos digitales, que es llamada SINTESIS DE VOZ. La síntesis de voz involucra la conversión de una secuencia comando o texto de entrada (palabras y oraciones), dentro de formas de onda de voz, usando algoritmos y previamente codificando datos de voz. El texto puede ser introducido por teclado, reconocimiento óptico de caracteres, o desde una base de datos previamente almacenada. En la figura 1.4.1. se muestra el diagrama a bloques de un sintetizador de voz.

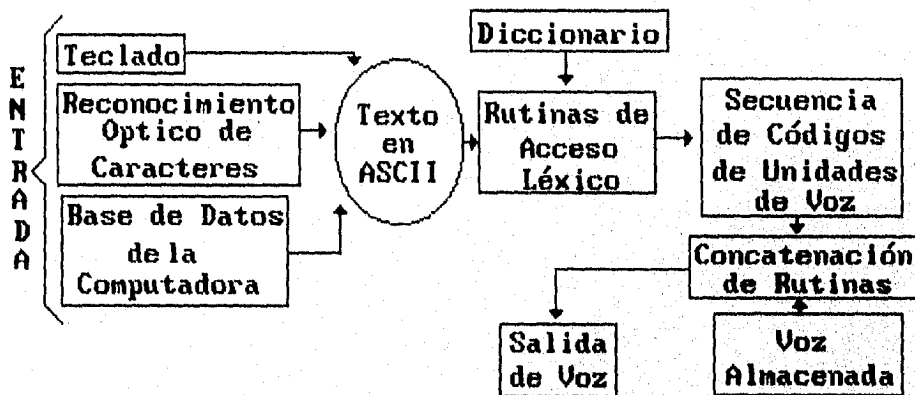


Fig.1.4.1. Diagrama a bloques de un sintetizador de voz.

La voz sintética fué generada primeramente en grandes computadoras, algunas veces interfaseada con un modelo del tracto vocal analógico, antes que los chips de procesamiento de señales de propósito especial fueran fabricados. Se puede decir que la mayoría de los sintetizadores de voz, usan un ancho de banda de 3Khz y 5Khz (aplicaciones telefónicas y para tener una calidad mayor, respectivamente). El uso de estas frecuencias se debe a que las vocales son percibidas muy bien con una frecuencia de 3Khz, debido a que pueden ser especificadas por los tres primeros "formants" o resonancias del tracto vocal (la frecuencia "formant" es inversamente proporcional a la longitud del tracto vocal, entonces en un adulto masculino ocurrirá con un promedio de 1000 Hz, y en un niño la tercera resonancia ocurre a los 300 Hz). La percepción de algunas consonantes es dañada ligeramente, si no se conserva este ancho de banda.

Los sintetizadores de voz comunes presentan un compromiso (trade-off) entre las demandas de maximización de calidad de la voz, aunadas a la minimización del espacio de memoria, complejidad del algoritmo y el tiempo de programación. Los sintetizadores de voz pueden ser caracterizados por el tamaño de las unidades de voz que se concatenan para producir la voz de salida, así como por el método usado para codificar, almacenar y sintetizar voz.

1.4.1. TAMAÑO DE LAS UNIDADES DE VOZ.

Podemos hablar de unidades largas o grandes y unidades pequeñas de voz. Las unidades largas, tales como frases u oraciones se concatenan en los sintetizadores más simples, en este tipo de unidades, se introduce alta calidad de voz a la salida (dependiendo del método de síntesis), pero necesita grandes cantidades de memoria para que todas las palabras o frases sean sintetizadas, después de que han sido habladas aisladamente. Para una voz sintética con máxima naturalidad, cada palabra o frase debe ser originalmente pronunciada con la apropiada sincronización y entonación, para todas las oraciones en las cuales serán usadas. Así, si una palabra ocurriera en varias estructuras sintéticas, la pronunciación será recordada y almacenada. Esto se logra concatenando las palabras originalmente habladas en aislamiento, lo que conduce a una pérdida de inteligibilidad y naturalidad. También existe una degradación en la calidad de la voz cuando se reduce la capacidad de la memoria.

Cuando nos referimos a unidades pequeñas de voz, hablamos de la unidad mínima de las palabras, llamada FONEMA (el lenguaje Español tiene 29 vocales y consonantes). Se puede generar la voz desde una secuencia de fonemas, el problema es que la pronunciación de un fonema en una palabra o frase es fuertemente dependiente del contexto en que se encuentra. La necesidad de calcular una apropiada entonación, da como resultado sintetizadores complejos y una salida de voz menos natural. Sintetizadores híbridos (usados en sistemas de investigación), concatenan inmediatamente el tamaño de unidades de voz almacenadas, utilizando la igualación de parámetros espectrales en los límites entre unidades.

1.4.2. PROPIEDADES BASICAS DE LA VOZ.

El proceso de producción de la voz, comunmente se modela como la respuesta de un filtro lineal en función del tiempo (el tracto vocal entre el glotis y los labios), como se muestra en al figura 1.4.2. En este modelo eléctrico, la voz es generada por una fuente de sonidos, la cual puede consistir de una excitación de una forma de onda de ruido, de una forma de onda de pulsos periódicos, o bien de una combinación de los dos. Como pulsos periódicos durante las regiones sonoras de voz, que es cuando el aire de los pulmones pasa entre las clausuras de las cuerdas vocales, es decir, cuando éstas estan vibrando. Como ruido durante regiones insonoras de voz, a causa de las turbulencias en puntos de constricción en el tracto vocal. El filtro acústico es el que simula el tracto vocal, la excitación es transformada a parámetros espectrales.

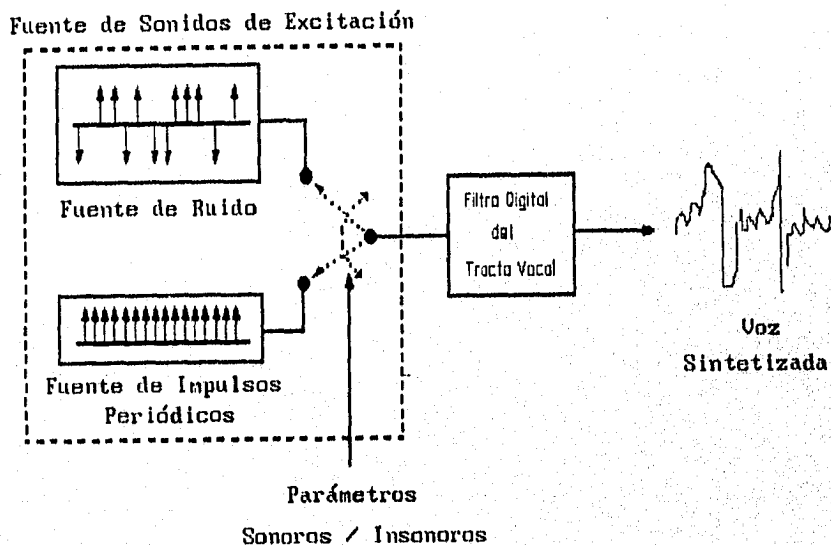


Fig.1.4.2. Modelo Eléctrico básico de producción de voz.

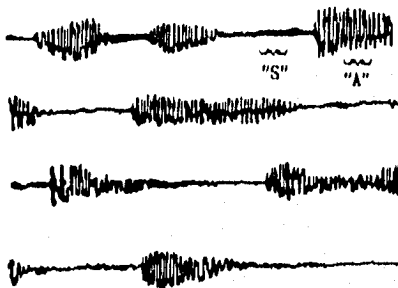
Considerables redundancias muestran las señales de voz, debido a los mecanismos físicos del tracto vocal humano y de la estructura inherente del lenguaje. Así mismo, a causa de los mecanismos físicos del oído, y debido a nuestra habilidad para percibir sonidos de voz. Las propiedades de producción y percepción de voz están resumidas en la tabla 2.1. Para observar las propiedades de producción de voz, tomaremos el ejemplo que se muestra en la figura 1.4.3a y b.

En la figura 1.4.3a, podemos observar períodos de tiempos largos y cortos de voz. En los tiempos largos vemos el comportamiento no estacionario, debido a la variación de la amplitud, y el comportamiento "sonoro/insonoro/silencioso", se ve variando con el tiempo (figura 1.4.3a'). En los tiempos pequeños (figura 1.4.3a'') se puede apreciar que es localmente estacionaria. El modelo espectral del tiempo pequeño de voz sonora e insonora, se muestra en la fig. 1.4.3b. Como se puede observar, en el segmento sonoro se forman formas de onda de estructura pseudoperiódica, mientras en los insonoros, tiene una apariencia de ruido.

La envolvente del espectro, está determinada por la respuesta a la frecuencia del filtro o del tracto vocal. Las resonancias del filtro, se observan como picos en la envolvente espectral, y es referida como formants. En el proceso de percepción de la voz, el oído actúa como un analizador espectral de Fourier o como un banco de filtros, en donde cada canal del banco de filtros tiene su propio control automático de ganancia y sus limitaciones del rango dinámico.

Debido a la construcción del oído, se produce el fenómeno de enmascaramiento auditivo, por lo que con señales que su espectro en frecuencia se encuentra fuera de la gama de los 20Hz a 20KHz, no son audibles. El oído igualmente es insensitivo a desfases, a distorsiones y ruidos que están fuera de la banda audible (ultrasónica o infrasónica). El umbral del ruido enmascarado se puede apreciar en la figura 1.4.3b. Tomando ventaja de las redundancias y restricciones en la producción y percepción de voz, es posible comprimir la razón de bits necesario para codificar y transmitir señales de voz. Las diversas técnicas de codificación hacen uso de estas ventajas.

Segmentos Largos de
Voz (600 mseg)



Segmentos Pequeños
de Voz (25 mseg)

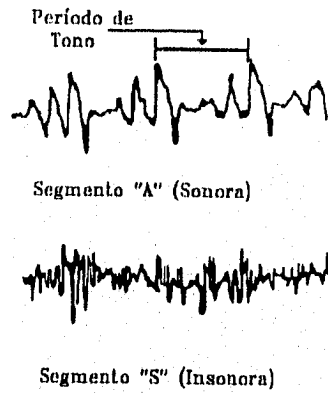


Fig.1.4.3.a Segmento de voz en tiempos largos y pequeños.

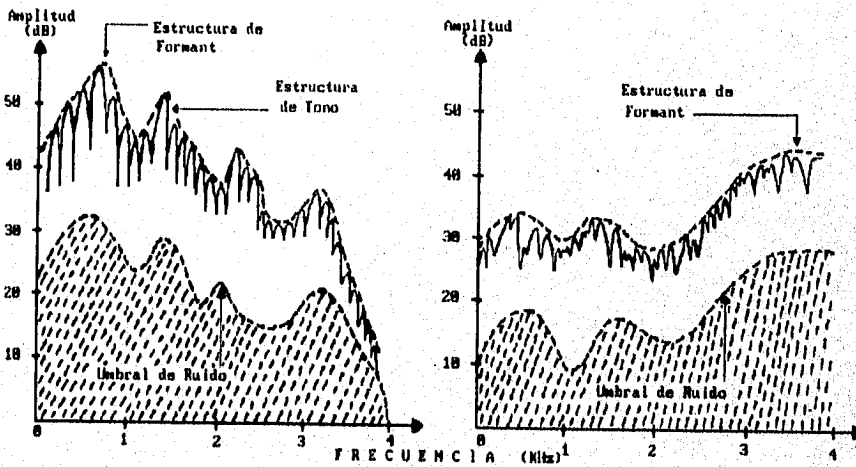


Fig.1.4.3.b Modelo espectral para voz sonora e insonora.

1.5. METODOS Y TECNICAS DE SINTESIS.

Los sintetizadores pueden ser clasificados dentro de dos clases:

- SISTEMAS TEXTO A VOZ.
- SISTEMAS DE RESPUESTA SONORA DE VOZ.

Los primeros sintetizan la voz constructivamente desde un texto, usando unidades pequeñas de voz, y procesando una lingüística extensiva. La segunda, produce la voz directamente de voz codificada previamente, utilizando técnicas de procesamiento de señales. También pueden ser clasificados los sintetizadores, de acuerdo a como parametrizan la voz para su almacenamiento y síntesis, los cuales pueden quedar de la siguientes formas :

CODIFICADORES DE FORMA DE ONDA.
CODIFICADORES DE LA FUENTE (VOCODERS) O PARAMETRICOS.
CODIFICADORES HIBRIDOS.

1.5.1 CODIFICACION DE FORMA DE ONDA.

La codificación de forma de onda, se caracteriza por que trata de reproducir exactamente la forma de onda de la señal original, cualquiera que sea la naturaleza de ésta (voz, música, datos en banda voz, etc.), y consiste básicamente en muestrear, cuantizar y codificar. la codificación de forma de onda, puede ser :

- A - EN EL DOMINIO DEL TIEMPO.
- B - EN EL DOMINIO DE LA FRECUENCIA.

A - TÉCNICAS EN EL DOMINIO DEL TIEMPO.

Las técnicas de codificación de forma de onda en el dominio del tiempo, se resume en la tabla 2.2. Este tipo de codificadores, muestrean periódicamente las formas de onda de las unidades de voz (el mínimo período se considera la razón de Nyquist), y se almacenan las muestras utilizando las técnicas digitales de la tabla. La mayoría de los codificadores en el dominio del tiempo, buscan minimizar el error o diferencia media cuadrática originada por el ruido cuantizado (debido a la representación de muestras de voz analógica, usando un número finito de niveles de amplitud), entre la voz original y la versión sintética, usando muestras codificadas.

En los esquemas de síntesis en el dominio del tiempo las cantidades de voz almacenadas en la memoria, son representaciones comprimidas de una forma de onda como función del tiempo. La acción del sintetizador, es simplemente desempaquetar dichas formas de onda que fueron almacenadas, para producir una señal de voz a la salida. Los circuitos VLSI para síntesis en el dominio del tiempo son simples, debido a que el "hardware" en esta técnica, no es necesario para convertir el dato almacenado del dominio de la frecuencia al dominio del tiempo.

La síntesis de voz en el dominio del tiempo es una tecnología joven, sin embargo, se han hecho rápidos progresos en la calidad de voz y en la razón de bit. El aspecto revolucionario de la técnica en el dominio del tiempo no es la implementación de "hardware", sino el análisis que permite que las formas de onda de voz puedan ser almacenadas en una forma altamente comprimida.

B.- TÉCNICAS EN EL DOMINIO DE LA FRECUENCIA.

Las técnicas de codificación en el dominio de la frecuencia, es la segunda categoría básica de codificación de forma de onda. Dentro de estas técnicas tenemos :

- SBC (Codificación en Sub-Bandas.)
- ATC (Codificación Adaptiva por Transformada.)

Ambos métodos proporcionan buena calidad de voz a una razón de bit de 16 Kbps. Los esquemas de síntesis en el dominio de la frecuencia, modelan la voz humana como una combinación de un pequeño número de fuentes de excitación (oscilaciones de las cuerdas vocales y turbulencia de aire), y una variedad más grande de estados de filtros de salida (representación de los estados de resonancia del tracto vocal). La compresión de datos se realiza, almacenando la excitación del tracto vocal y se filtran los parámetros, en lugar de la forma de onda original.

El "hardware" para síntesis en el dominio de la frecuencia, contiene circuitos integrados relativamente complejos (muchos polo analógicos o filtros digitales), que transforman los parámetros en el dominio de la frecuencia dentro de una señal de audio en el dominio del tiempo, con un espectro en frecuencia en función del tiempo, el cual aproxima el espectro de la forma de onda original.

Las técnicas en el dominio del tiempo y en el dominio de la frecuencia, serán tratados más ampliamente en los capítulos subsecuentes.

1.5.2. CODIFICACION DE LA FUENTE (VOCODING) O PARAMÉTRICA.

Los "vocoders" sintetizan la voz en base a un modelo que simula el aparato productor de voz humana, e intenta conseguir el mismo espectro de la voz original, en lugar de recrear la forma de onda muestra por muestra. Los canales "vocoder" están entre los tipos más primitivos históricamente de técnicas de compresión de voz. Los codificadores de la fuente, almacenan un número limitado de parámetros que son relevantes perceptualmente de las unidades de voz, para reducir el espacio de voz. La voz estará dividida dentro de pequeños marcos (10 o 30 mseg.) y se determinan los parámetros describiendo el espectro y la periodicidad de la voz dentro del marco.

Se puede hacer un análisis tosco de la amplitud del espectro, aprovechando que el oído es relativamente insensible a la distorsión de fase. En la elección del tamaño del marco, hay un compromiso entre almacenaje y calidad de voz, debido a que si los marcos son más pequeños, entonces requerirán más almacenaje. En la tabla 2.3, se muestran las técnicas de codificación de la fuente.

El banco de filtros fué la primera construcción realizada para el analizador, y el análisis espectral de este banco de filtros se puede lograr utilizando la transformada discreta de Fourier. Este banco de filtros, fué sustituido por el canal "vocoder" en el analizador. Los parámetros de excitación obtenidos por el analizador, son usados para generar una excitación consistente de pulsos espaciados por el período de tono (pitch) para la voz sonora y una secuencia de ruido para los sonidos insonoros.

En la función de autocorrelación de tiempo pequeño de la onda de voz, la longitud de la ventana o marco de autocorrelación exhibirá picos en los múltiplos de los períodos de tono. Consecuentemente con estos picos, la frecuencia fundamental puede ser determinada. Por otro lado, la ausencia de picos sirve para identificar intervalos insonoros o de silencio.

La forma de onda de la voz, consiste de una convolución de la función de excitación y la respuesta al impulso del tracto vocal. Para la desconvolución de la voz se puede adaptar el filtrado homomórfico, que es la no linealidad general de las técnicas de filtración. Para sistemas de análisis y síntesis de voz basados en la filtración homomórfica, el valor del "cepstral" de tiempo pequeño es usado como parámetro para representar el tracto vocal o su equivalente, la envolvente espectral de la voz. Los parámetros de excitación son tomados de la porción de tiempos largos del "cepstrum" en los múltiplos del tono periódico, y para voz insonora no ocurren tales picos.

En el sintetizador homomórfico, la entrada consiste de un valor "cepstral" de tiempo pequeño y los parámetros de excitación. En el sintetizador, los valores del "cepstral" de tiempo pequeño son convertidos a una aproximación de la respuesta al impulso del tracto vocal. Los parámetros de excitación, son usados para generar una función de excitación, la cual es convolucionada con la aproximación de la respuesta al impulso, obtenida en el "cepstral", con lo que resulta la voz a la salida.

SINTETIZADORES DE FORMANT.

El sintetizador paramétrico que se ha empleado más frecuentemente utiliza codificación de "formant" y anchos de banda para representar el espectro almacenado. Esta técnica, separa la forma de onda dentro de una componente de excitación y una representación del filtro de tracto vocal. La excitación es un tren de impulsos periódicos para voz sonora y ruido blanco para voz insonora. El tracto vocal es usualmente modelado como una serie de resonadores digitales de segundo orden, cada uno modelando un "formant" o el efecto de la cuerda vocal. Esto aproxima bien el espectro de voz para vocales, y puede ser controlado con un parámetro de amplitud.

Los sonidos insonoros tienen la fuente de excitación dentro de la cavidad vocal, excitando primeramente el tracto vocal entre los labios y la fuente, con lo que el tubo acústico está recortado y la mayoría de la energía está a frecuencias muy altas. Para estos sonidos, es comúnmente usado un banco de resonadores en paralelo, ya que cada filtro puede ser individualmente controlado por amplitud. Los sonidos nasales generalmente necesitan un resonador extra, debido a que la trayectoria acústica que incluye la cavidad nasal es más grande que para las vocales, por lo que se incrementa el número de resonancias en el ancho de banda de la voz. La carencia de variación en tiempo en los anchos de banda, es más perceptible en este tipo de sonidos, que están sujetos a grandes pérdidas en la cavidad nasal, por lo que resulta ser "formants" más amplios.

En los sonidos nasales, son encontrados ceros espectrales a frecuencias donde la cavidad vocal presenta impedancia cero para la presión de las ondas saliendo de la faringe. Los ceros pueden directamente ser introducidos utilizando un resonador inverso de segundo orden, dentro del filtro que simula el tracto vocal. El resonador extra y el resonado inverso, se pueden considerar como un par de polos-ceros, con efectos de cancelación, excepto cuando sonidos nasales son simulados. Sin embargo, la localización de polos y ceros automáticamente en voz natural es una labor difícil, ya que una modelación cero con baja calidad es perceptible en estos sonidos.

CODIFICACION DE PREDICION LINEAL.

Otro sintetizador ampliamente usado es el LPC, que al igual que el de "formant", separa la forma de onda dentro de una componente de excitación y una representación del tracto vocal. Sin embargo, estos sintetizadores son más simples conceptualmente, todos los efectos espectrales están incluidos en los parámetros LPC, los cuales son calculados automáticamente en el análisis de la voz original. La mayoría de los sistemas LPC son representados por el filtro del tracto vocal "solamente polo" (all pole). A causa de que el modelo LPC, no puede separar los efectos de la radiación de los labios, ni los efectos de las cuerdas vocales desde la respuesta al tracto vocal, existen generalmente un máximo de cuatro "formants" excitados abajo de los 4Khz, y cada uno está bien modelado por un resonador de segundo orden, es decir, un par de polos complejos en el plano espectral Z. Para describir la filtración entre las cuerdas vocales y los labios, son suficientes 8 polos.

El filtro digital solamente polo de síntesis predice una muestra de voz basada sobre una combinación lineal de N muestras previas. El problema de los sistemas LPC solamente polos, es que la excitación glotal introduce ceros dentro del espectro cuando se producen sonidos nasales, debido a que estos son radiados desde más de una parte del tracto vocal o cuando existe una ramificación entre la fuente de excitación y la salida de voz. Los ceros generalmente causan modificaciones en resonancias adyacentes, especialmente en sus anchos de banda.

Si el filtro LPC es excitado por impulsos periódicos para voz sonora, el tracto vocal es excitado una vez por el período de tono (pitch) en el instante específico de tiempo, y la duración de la excitación, no es mucho más grande que un período de muestra. Excitando el filtro de síntesis con un impulso solamente, causa una calidad menor de voz. Para reducir la razón de bit en el almacenamiento, puede ser usada una versión residual de la señal, la que podemos obtener si la voz sonora es pasada a través de un filtro cuya respuesta es el inverso del filtro de síntesis LPC, la salida de éste es la señal residual, lo que generalmente contiene impulsos fuertes en los tiempos cuando las cuerdas vocales están cerradas. Sin embargo, la distribución de la energía residual es significativa en todas partes del período de tono.

Los sintetizadores LPC, son más populares que los otros "vocoders", debido a su algoritmo sencillo, produciendo una representación espectral apropiada para una codificación económica. En los "vocoders", se requiere a priori conocer el mecanismo por medio del cual se genera la voz, debido a que éstos simulan la voz humana. Además, los conocimientos de la estructura fonética y del lenguaje son importantes para aplicarlos en los parámetros del "vocoder", y de esta forma conseguir una voz sintética más natural.

Todas las técnicas de codificación de la fuente se explicaron ligeramente, pero las más importantes comercialmente y más frecuentemente utilizadas son las LPC y la de "formants".

1.5.3. CODIFICACION HIBRIDA.

Esta tercera clase de codificación, como su nombre lo indica, combina las características de codificación de forma de onda y la de codificación de la fuente, en un intento de llenar el vacío que existe entre el funcionamiento de ambos métodos. Dentro de esta clase de codificación tenemos :

- Vocoder Excitado por Voz (VE-V)
- Escalamiento Armónico.

La primera técnica, utiliza métodos de codificación de forma de onda en la primera porción de la banda de voz, de 0 - 800 Hz, y el resto de la banda de voz entre 800 - 4000 Hz, utiliza la codificación de la fuente. La excitación de la fuente para la parte de la codificación de la fuente, la banda de voz se obtiene por una distorsión no lineal de la banda base de la forma de onda codificada. En la segunda técnica, la señal de voz es comprimida paramétricamente en el ancho de banda y el rango de muestreo. Esta señal comprimida puede ser codificada digitalmente por métodos de codificación de forma de onda para obtener una razón de bit reducida. La señal comprimida, puede utilizarse para "acelerar" o "retardar" la voz.

El método de escalamiento armónico se enfoca primeramente sobre redundancias de la voz, debido a la estructura de tono y a que es estacionaria localmente, por lo que trabaja mejor con técnicas de codificación que no lo hacen depender sobre estas propiedades. Este método se ha realizado en el dominio del tiempo y también en el dominio de la frecuencia.

El escalamiento armónico en el dominio de la frecuencia es una técnica que comprime o expande el espaciado de las armónicas de tono. En el dominio del tiempo, es un método que elimina o comprime, o por el contrario inserta o expande los períodos de tono en un tono de manera sincrónicamente interpolados. El área de codificación híbrida está activamente estudiada, para que la codificación de forma de onda y la codificación de la fuente se reduzca dentro de un cuadro más unificado.

1.5.4. SISTEMAS DE TEXTO ILIMITADO (TEXTO A VOZ) Y SISTEMAS DE TEXTO LIMITADO (RESPUESTA SONORA).

Como se dijo anteriormente, también los sintetizadores son clasificados como :

- Sintetizadores de Texto a Voz.
- Sintetizadores de Respuesta Sonora.

Los sistemas de texto a voz tienen la necesidad de un procesador lingüístico para convertir el texto dentro de una forma apropiada para el acceso de unidades de voz almacenadas. Por lo que deben trasladar las oraciones de entrada dentro de una secuencia de códigos para extraer las apropiadas unidades almacenadas, y determinar los parámetros de entonación del texto. Para poder realizar el traslado, se maneja por medio del conjunto de reglas del lenguaje para convertir una serie de letras dentro de una secuencia de fonemas.

Las reglas examinan el contexto de cada letra para determinar como deben ser pronunciadas. Un lenguaje como el Español necesita pocas reglas pues se puede escribir con buena ortografía siguiendo cercanamente la fonética, ya que cada letra tiene solamente una pronunciación. Los sistemas de respuesta sonora son ampliamente utilizados en aplicaciones como teléfonos con asistencia automática de directorio, juguetes que hablan, sistemas de protección de máquinas y automóviles.

Estos sistemas codifican esencialmente la voz del usuario con eficientes métodos de almacenamiento, por lo tanto, están limitados por las combinaciones de las unidades originales de voz, usualmente, con las entonaciones originales.

1.6. CALIDAD DE VOZ.

La calidad de la voz es un factor importante de considerar en los sintetizadores; se puede mostrar el "estado del arte" de síntesis en términos de calidad de voz (el rango 0 - 1, es una medida hipotéticamente simplificada, donde el cero implica una pobre inteligibilidad, y el uno implica una calidad que es indistinguible esencialmente de la entrada original par aun a señal en el ancho de banda telefónica) como una función de la razón de bit, lo cual se muestra en la figura 1.6.1.

En esta figura, se muestran los tres métodos de codificación, de forma de onda, de modelación de la fuente y la híbrida. La calidad se muestra como una razón de bit, debido a que la disminución de la razón de bit es importante para que las técnicas de codificación ocupen un almacenamiento menor en la memoria.

Los métodos de codificación de forma de onda, sin hacer uso de las propiedades de la tabla 2.1, se puede obtener buena calidad de voz a una razón de bit sobre los 64 Kbps. Si se toma ventaja de la amplitud no estacionaria, se puede lograr una buena calidad con esquemas simples de codificación de forma de onda con la misma razón de bit. Tomando ventaja de las características espectrales estáticas y dinámicas, se puede lograr una buena calidad a una razón de bit en los rangos de 24 - 32 Kbps. Una calidad aceptable puede lograrse con una razón de bit bajo los 10 Kbps si se incluyen las propiedades de tono y ruido. La calidad de los métodos de forma de onda debajo del rango de los 10 Kbps, caen rápidamente como una función de la razón de bit y de las propiedades de la voz.

La razón de bit para una modelación de la fuente, puede ser reducida al orden de los 2 Kbps, obteniéndose una voz como máquina. La calidad de esta técnica está limitada por el modelo de la señal "vocoder" y por las propiedades del lenguaje y la estructura fonética. Las técnicas de codificación que han estado evolucionando son los métodos híbridos o combinados, los cuales pueden trabajar de 4 a 10 kbps, donde al utilizar las ventajas de los métodos de forma de onda y de la modelación de la fuente, obtiene buena calidad de voz con este rango de razón de bit.

En todos los métodos de codificación, se pueden conseguir reducciones adicionales en la razón de bit en grandes períodos de tiempo, si se corta el codificador en las regiones de silencio, es decir, cuando el orador no está hablando. Realizando esto, conduce a técnicas de rango variable de codificación, la razón promedio de bit dependerá de la actividad del orador y el porcentaje de tiempo que esté hablando.

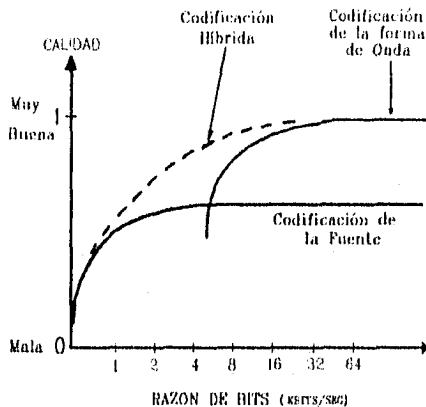


Fig.1.6.1. Calidad vs Razón de bit en diferentes técnicas

1.7. ESTADO ACTUAL DE LA TECNOLOGIA PARA IMPLEMENTAR SINTESIS DE VOZ.

El procesamiento digital de señales de voz, se conoce desde hace poco más de 25 años con una gran actividad y con un amplio rango de aplicaciones en las comunicaciones y en desarrollo de las máquinas lectoras y de respuesta hablada. Sin embargo, actualmente el gran auge tomado es porque la tecnología tiene la factibilidad de realizar algoritmos complejos en forma económica. Los avances logrados en microelectrónica permite implementar algoritmos en unos pocos circuitos integrados, por lo que, todos los tipos de representaciones de las señales se han desarrollado, pudiéndose implementar en tiempo real y a un costo razonable. En la fig.1.7.1., se muestra como se han incrementado y desarrollado las técnicas de codificación de voz, que se han podido implementar en forma económica en el transcurso de los últimos años.

Por las perspectivas que se presentan de un gran mercado para esta área, los fabricantes de semiconductores a nivel mundial se han avocado al desarrollo de procesadores de señales cada vez más rápidos y poderosos. El avance de la tecnología VLSI proporciona un gran camino a generaciones de procesadores, y ha sido tan rápida que la primera generación inicia en 1979 con el 2920 de INTEL, en 1982 se da la segunda generación con el TMS32010 de TEXAS INSTRUMENT, y la tercera generación llega con el TMS3477 de TEXAS INSTRUMENT. Esto puede observarse gráficamente en la fig.1.7.2.

La tecnología VLSI ha hecho posible el uso de chips en los años recientes para aplicaciones digitales de la voz. En el área de codificación de forma de onda, se han introducido chips para conversión completa A/D y D/A en PCM ley u que incluye filtrado anti-aliasing, chips para ADM, ambos con interés de aplicación en telefonía y también con creciente interés en esta área a 32 kbps chips ADPCM. El el área de "vocoders" también están disponibles los sintetizadores en chips, como el utilizado en juguetes educativos con respuesta sonora.

El área de la tecnología VLSI con fuerte impacto en voz digital, es el desarrollo de microprocesadores de alta velocidad y circuitos integrados programables, con los que es posible implementar algoritmos como el ADPCM y el SBC (Codificación en Sub-Bandas). Para algoritmos de alta complejidad como el ATC (Codificación Adaptiva por Transformada), APC (Codificación Predictiva Adaptiva) y Analizadores "Vocoder", son realizados en tiempo real usando procesadores o "hardware" de propósito especial. Ejemplos de este tipo de circuitos integrados de propósito especial de procesamiento de voz son :

- MSM5218 (ADPCM Analizador y Sintetizador) de OKI.
- TMS5220 (Síntesis por LPC) de TEXAS INSTRUMENT.
- D2911 (Codec PCM ley A) de INTEL.
- MC3518 (CVSD) de MOTOROLA.
- TMS3477 (CVSD) de TEXAS INSTRUMENT.

y su número tiende a aumentar en el mercado de semiconductores.

- PCM Modulación por Codificación de Pulsos.
- DM Modulación Delta,
- CH-V Vocoder de Canal.
- DPCM PCM Diferencial.
- ADPCM PCM Diferencial Adaptivo.
- CVSD Deltamodulación de Pendiente Variable Cuntínua.
- LPC Código de Predicción Lineal.
- FOR-V Vocoder de Formants.
- SBC Codificación en Sub-Bandas.
- VE-V Vocoder Excitado por Voz.
- TES Voz Codificada en Tiempo.
- APC Código de Predicción Adaptiva.
- ATC Codificación Adaptiva por Transformada.
- VQ Vector de Cuantización.
- THDS Escalamiento de Armónicas en el Dominio del Tiempo.

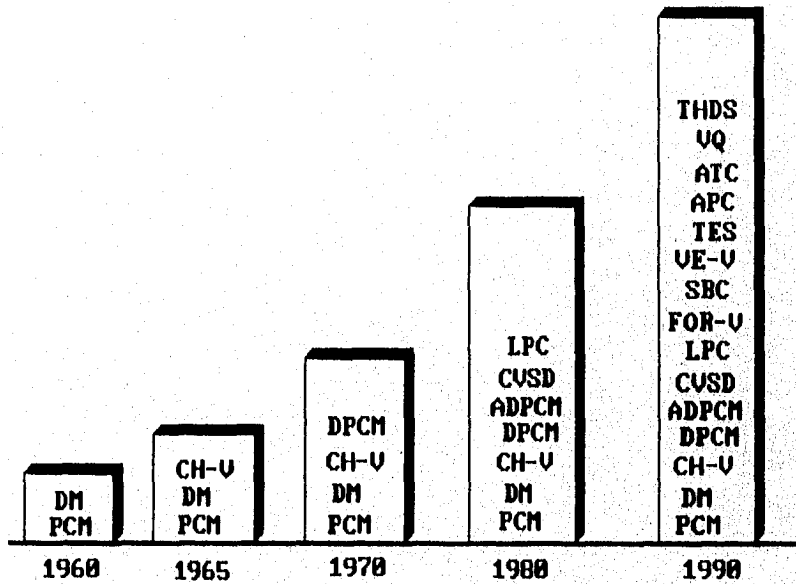


Fig.1.7.1. Crecimiento de las Técnicas de Codificación.

PROCESADOR DIGITAL DE SEÑALES	1980	81	82	83	84	85	86	87	88	89
GENERACION	1a	1a	2a	2a	3a	3a	3a	3a	3a	3a
NEC7720	X		Y							
BELL-DSP	X					X*				
AMD29500						Y				
TI-320			X	Y		X*				
FUJITSU 8764				X						
HITACHI HD61810			X		X					
NEC7282					X					
IT-UDPI-01					X					
STO-DSP-28				X						
NATIONAL DSP							X	Y		
NEC77230							X	Y		
OKI MSM 5218RS							X			
TEXAS TMS3477							X			

- X = ANUNCIO INICIAL
 Y = SEGURO DE USO
 Z = ABSOLETO
 X* = CAMBIA SU DISEÑO DE "NMOS" A "CMOS"

Fig. 1.7.2 Estado actual de los procesadores Digitales de Señales.

1.8. COMPROMISO ENTRE COMPLEJIDAD Y FUNCIONAMIENTO.

En el diseño de un codificador de voz, existen tres clases de compromisos principales:

- CALIDAD TOTAL.
- COSTO O COMPLEJIDAD.
- RAZON DE BIT REQUERIDO.

Estos compromisos se ilustran en la fig.1.8.1., para algunas realizaciones de "hardware". La escala vertical corresponde a la calidad de la voz, medida de 1 a 5 basas sobre un tanteo de opinión media para el juicio de los oyentes. La escala horizontal ofrece una medida brusca de complejidad "hardware".

Las diferentes curvas representan la razón de bit y los puntos sobre las curvas corresponden a la realización "hardware" que ha sido implementada en el laboratorio. Por tanto, de la figura es posible obtener una valorización relativamente buena de los compromisos de calidad, complejidad y razón de bit.

Tabla 2.1.- PROPIEDADES BÁSICAS DE PRODUCCIÓN Y PERCEPCIÓN DE VOZ.

A.- PRODUCCIÓN DE VOZ
1.- NO ESTACIONARIA EN TIEMPOS LARGOS
- Amplitud no estacionaria.
- Regiones [Sonoras/Insonoras/Silencios].
- Estructura del lenguaje y la fonética.
2.- ESTACIONARIA EN TIEMPOS CORTOS
- Estructura Formant (correlación en tiempos pequeños)
- Estructura de Tono "Pitch" (sonoros).
- Estructura de Ruidos (insonoros).
B.- PERCEPCIÓN DE VOZ
- Rango dinámico espectral local.
Enmascaramiento o disimulamiento auditivo.

Tabla 2.2.- TÉCNICAS DIGITALES DE CODIFICACIÓN EN EL DOMINIO DEL TIEMPO.

- PCM	Modulación por Codificación de Pulsos (Lineal y Logarítmica).
- DPCM	Modulación por Codificación de Pulsos Diferencial.
- ADPCM	Modulación por Codificación de Pulsos Diferencial Adaptivo.
- DM	Modulación Delta.
- ADM	Modulación Delta Adaptiva.
- CVSD	Modulación delta con Pendiente que Varía Continuamente.
- APC	Codificación por Predicción Adaptiva.

Tabla 2.3.- TÉCNICAS DE CODIFICACIÓN DE LA FUENTE.

- Banco de Filtros.
- Vocoder de Canal.
- Modelo Homomórfico.
- Autocorrelación.
- Modelo de Formant.
- Código ded Predicción Lineal (LPC).

CAPITULO 2

FUNDAMENTOS TEORICOS

INTRODUCCIÓN

Para el diseño de todo sistema se requiere contar con las bases teóricas que permitan establecer las especificaciones de funcionamiento, comparar con otros dispositivos similares, o en el mejor de los casos desarrollar nuevas fórmulas para la teoría en estudio.

En este capítulo se desglosan los fundamentos teóricos de la comunicación digital en general. En las primeras cuatro secciones se estudian los teoremas y resultados del PCM, como son: el teorema de muestreo, la cuantización, compansión, etc.; así como la descripción de un sistema telefónico real de PCM, el T1 de la Bell System, las características del ruido y además el diseño de filtros.

Se intenta desarrollar la mayor parte de las fórmulas básicas, sin embargo, por lo engorroso que resultaría seguir paso a paso las demostraciones, algunas veces se preferirá indicar el procedimiento que lleva a la expresión de los resultados y apreciar estos en sus detalles más significantes.

2.1 EL TEOREMA DE MUESTREO DE NYQUIST

Este teorema juega un papel importante en las comunicaciones modernas pues gracias a él se ha iniciado una práctica muy común y bastante eficiente, el muestreo a intervalos uniformes de una señal continua limitada en banda que contiene en sus valores discretos toda la información de esta señal. Entiéndase por señal limitada en banda aquella que no contiene componentes espectrales mayores a una cierta frecuencia de corte.

En el primer capítulo se mencionó a grandes rasgos el teorema de muestreo uniforme que se refiere a la especificación de una señal dada mediante muestras cuyas tomadas a intervalos uniformes de f_m segundos, correspondiendo a esta sección demostrar el porqué tomar a esta velocidad mínima el muestreo y no otra.

2.1.1. MUESTREO IDEAL

Sea $f(t)$ una señal limitada en banda que no contenga componentes espectrales mayores de f_m Hz, es decir que $F(\omega)$ la transformada de Fourier de $f(t)$ es cero para $\omega > \omega_m$. Supongamos que multiplicamos $f(t)$ por una función de impulsos periódicos de altura unitaria $\delta_T(t)$, la función producto es una sucesión de impulsos localizados a intervalos regulares de T segundos con altura correspondiente al valor de $f(t)$ en el instante de aparición de éstos. Es evidente que esta señal discreta representa las muestras instantáneas de la señal cada intervalo T de tiempo. Véase Fig.2.1.1.1.a,c,e.

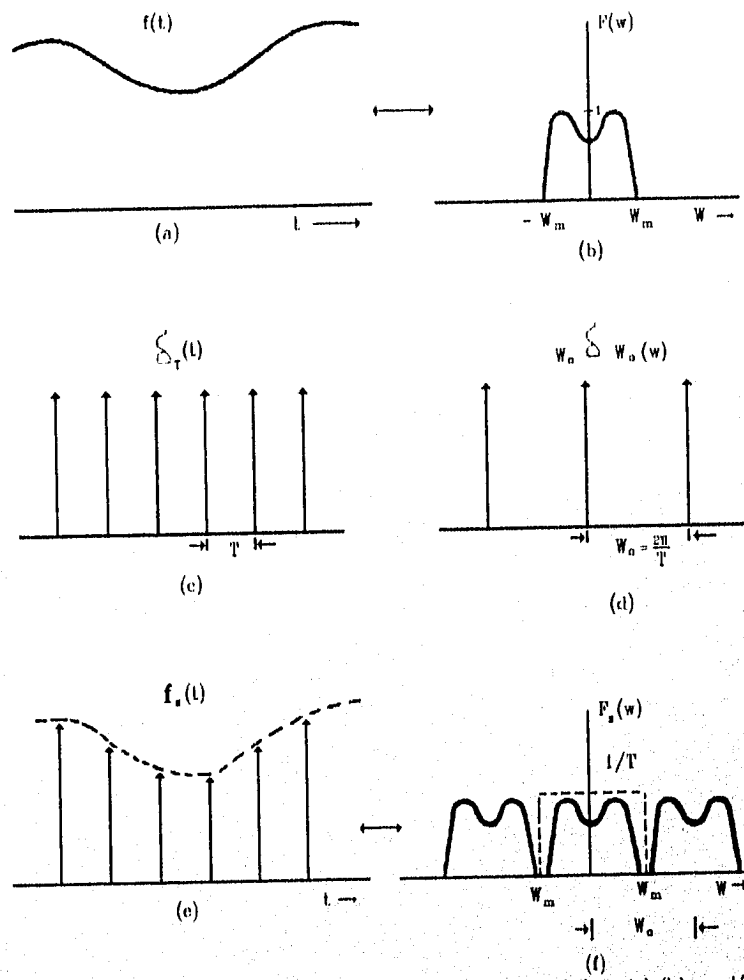


Fig. 2.1.1. Representación en tiempo y frecuencia de las señales: (a),(b) analógica; (c),(d) tren de impulsos; (e),(f) muestreada.

Representemos la señal muestreada por :

$$f_s(t) = f(t) \delta_T(t) \quad (2.1.1.1)$$

En los incisos (b) y (d) de la Fig.2.1.1.1 se muestran las transformadas de Fourier correspondientes a la señal limitada en banda y al tren uniforme de funciones impulso $\delta_T(t)$.

Para obtener la transformada de Fourier de $f_s(t)$ en la ecuación 2.1.1.1, aplicaremos el teorema de convolución en la frecuencia.

Por lo que:

$$F_s(\omega) = \frac{1}{2\pi} (F(\omega) * \omega_o \delta_{\omega_o}(\omega)) \quad (2.1.1.2)$$

$$\omega_o = \frac{2\pi}{T}$$

$$F_s(\omega) = \frac{1}{T} (F(\omega) * \delta_{\omega_o}(\omega)) = \frac{1}{T} (F(\omega) * \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_o))$$

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_o) \quad (2.1.1.3)$$

Esto implica que el espectro en frecuencia de la señal muestreada es el mismo de la señal limitada en banda repetido cada ω_o radianes por segundo. Obsérvese en la Fig.2.1.1.1.f que para poder recuperar la señal $f(t)$ es necesario hacer pasar $f_s(t)$ por un filtro paso bajo con frecuencia de corte f_m Hz.

Además es importante lo siguiente: para que sea factible recuperar la señal original es necesario que los espectros repetidos no se traslapen, condición que se cumple si :

$$\omega_o \geq 2\omega_m$$

o en términos del período :

$$\frac{2\pi}{T} \geq 2 (2\pi f_m)$$

$$T \leq \frac{1}{2f_m} \quad (2.1.1.4)$$

que es la condición de muestreo establecida por Nyquist, base primordial de todas las formas de transmisión por modulación de pulsos.

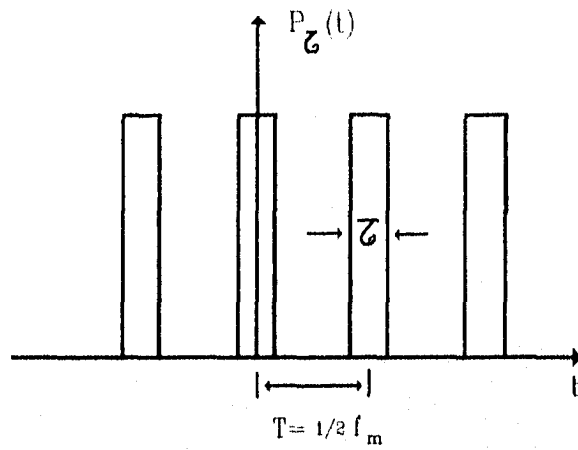
2.1.2 MUESTREO REAL

En la práctica no se muestrea por impulsos ideales, sino por pulsos de pequeña duración, lo que significa que el muestreo no es instantáneo sino que ocurre en tiempos finitos (Fig.2.1.2.1a). Por un proceso similar al anterior se demuestra que la señal muestreada tendrá la siguiente función espectral.

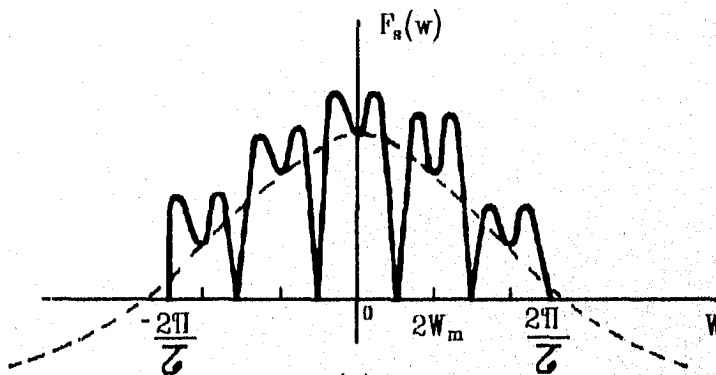
$$F_s(\omega) = \frac{\tau}{T} \sum_{n=-\infty}^{\infty} Sa(n\pi\omega_m) F(\omega - 2n\omega_m) \quad (2.1.2.1)$$

para el caso de $\omega_o = 2\omega_m$.

El segundo miembro de la ec.2.1.2.1 representa el espectro de $F(\omega)$ repetido cada $2\omega_m$ radianes por segundo con un factor de variación en su amplitud de $Sa(n\pi\omega_m)$ (Fig.2.1.2.1.b). Es importante darse cuenta que el ancho de banda necesario para transmitir una señal muestreada por impulsos es infinito, mientras que para el muestreo por pulsos tiene un valor finito, de modo que el contenido de energía de $F_s(\omega)$ es despreciable en frecuencias superiores.



(a)



(b)

Fig. 2.1.2.1. Muestreo real. (a) pulsos de muestreo, (b) espectro de la señal muestreada.

2.1.3. RECONSTRUCCION DE LA SEÑAL ANALÓGICA

El proceso analítico que llevaremos a cabo se hará en base a considerar muestras por impulsos para evitar entrar a cálculos más complejos y menos ilustrativos, siendo los resultados obtenidos bastante similares a cuando el muestreo se realiza en forma real.

Consideremos que muestreamos con la rapidez mínima, es decir:

$$T = \frac{1}{2f_m}$$

$$\omega_o = \frac{2\pi}{T} = 2\omega_m$$

y sustituyendo en la ec. 2.1.1.3

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - 2n\omega_m) \quad (2.1.3.1)$$

Es evidente que podríamos recuperar $F(\omega)$ si multiplicamos esta ecuación por una función pulso rectangular $R_{2\omega_m}$. Lo que conduce a :

$$F_s(\omega)R_{2\omega_m} = \frac{1}{T} F(\omega)$$

por lo que

$$F(\omega) = TF_s(\omega) R_{2\omega_m}(\omega) \quad (2.1.3.2)$$

Donde $TR_{2\omega_m}(\omega)$ representa al filtro paso bajo con una frecuencia de corte ω_m y

una ganancia $T = \frac{1}{2f_m}$. Aplicando el teorema de convolución en el tiempo a la

ec.2.1.3.2.

$$f(t) = Tf_s(t) * \frac{\omega_m}{\Pi} S_a(\omega_m t)$$

$$f(t) = f_s(t) * Sa(\omega_m t)$$

siendo

$$f_s(t) = \sum_{n=-\infty}^{\infty} f_n \delta(t-nT)$$

donde f_n es la n ésima muestra de $f(t)$. Prosiguiendo

$$f(t) = \sum_{n=-\infty}^{\infty} f_n \delta(t-nT) * Sa(\omega_m t)$$

$$f(t) = \sum_{n=-\infty}^{\infty} f_n Sa(\omega_m(t-nT))$$

$$f(t) = \sum_{n=-\infty}^{\infty} f_n \text{Sa}(\omega_m t - n\pi) \quad (2.1.3.3)$$

De aquí se observa que es posible reconstruir $f(t)$ sumando los productos de cada una de las n muestras con la función muestreo. En la Fig.2.1.3.1 se ilustra este proceso mostrándose únicamente las contribuciones de tres muestras por simplicidad.

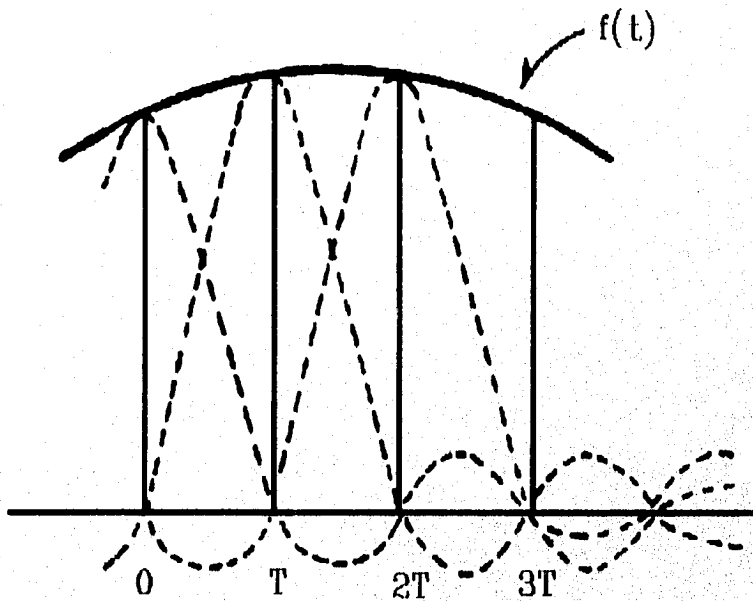


Fig. 2.1.3.1 Reconstrucción de la señal analógica.

2.2. REQUERIMIENTOS DE CODIFICACIÓN

Una vez que se obtiene la muestra, es necesario transformarla por métodos adecuados a un código que permita su transmisión con la mayor eficiencia. No siempre se logra esto debido a factores inherentes al sistema tales como el ruido o requerimientos de ancho de banda.

No es posible evitar un ruido de cuantización que obstaculiza reconstruir tal cual la señal analógica en el receptor, ya que la cuantización en sí trae la incertidumbre de ciertos niveles de la señal que son ajustados al nivel cercano más próximo. Una medida de como afecta el ruido al sistema es la relación señal a ruido de cuantización de la que hablaremos ampliamente en la sección 2.3. Gran número de bits trae consigo un incremento en el ancho de banda, aumentando aún más cuando se trabaja por división en el tiempo para acomodar varios canales de datos, es pues un dilema el número de niveles de cuantización versus ancho de banda.

2.2.1. NIVELES DE CUANTIZACIÓN

En las variaciones de amplitud existe la necesidad de valuarlas a ciertos niveles, tratando de minimizar la pérdida de información no eliminable de la cuantización. La cantidad de niveles dependerá de la precisión requerida; así en un sistema PCM donde sea necesario transmitir la altura de un líquido en un depósito industrial hacia una central de cómputo, en el cual interese únicamente registros de cambio de 1 m en excursiones que van de 0 a 15 m seguramente se utilizarían 16 niveles, mientras que si se requiriese conocer los incrementos de altura de 0.5 m es evidente el uso de 32 niveles de cuantización.

Un código puede realizar un cierto número de combinaciones, resultado de las diferentes maneras en que los pulsos puedan acomodarse uno con respecto a otro, de esta forma, siendo N el número de niveles de cuantización se tiene

$$N = m^n \quad (2.2.1.1)$$

donde m son los niveles de altura que pueden tomar los pulsos y n el número de estos. Así para m=2 obtenemos el código binario:

$$N = 2^n$$

o $n = \log_2 N \quad (2.2.1.2)$

Es pues n función logarítmica de los niveles de cuantización (Tabla 2.2.1.1).

N	n
2	1
4	2
8	3
16	4
32	5
64	6
128	7
· ·	·
· ·	·

Tabla 2.2.1.1

2.2.2. ANCHO DE BANDA

El incremento del ancho de banda es apreciable en cuanto existe un número n de bits que han de ser colocados en un "time slot" es inherente a cada muestra, por lo que al disponer de una señal limitada en banda a 4 KHz, como la voz, tendremos al aplicar el teorema de Nyquist una señal discreta con una frecuencia de 8Khz; si cuantizamos con 256 niveles ($n = 8$), el ancho de banda PCM resultante es de $8 \times 8 = 64$ KHz. Es evidente pues, que :

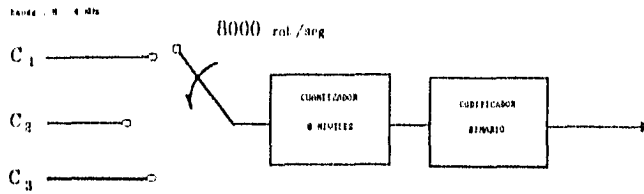
$$f_{PCM} = 2f_m n \quad (2.2.2.1)$$

En caso de multiplex por división en el tiempo, la ecuación anterior resultará afectada por un factor igual a l número de canales C_t que en tren en la transmisión, o sea :

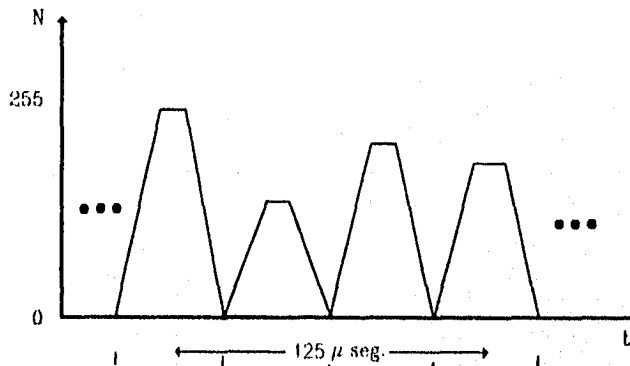
$$f_{PCM} = 2f_m n C_t \quad (2.2.2.2)$$

En la Fig.2.2.2.1 se muestra la aplicación de la ec.2.2.2.2 para 3 canales del mismo ancho de banda, suponiendo además que cada muestra ocupa el 100% de su "time

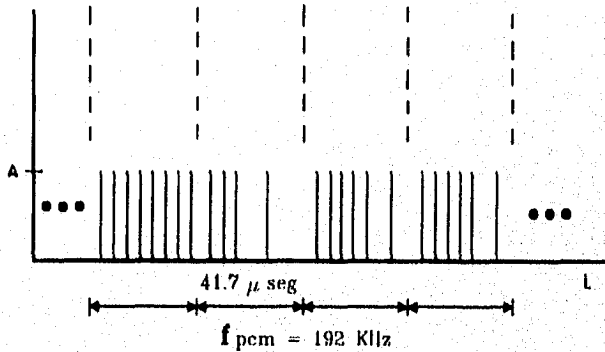
s | o t "



(a)



(b)



(c)

Fig. 2.2.2.1. Multiplex y conversión analógica - digital.

(a) diagrama a bloques.

(b) salida del cuantizador.

(c) salida del codificador.

2.3. RUIDO DE CUANTIZACIÓN

Este ruido es inherente a toda modulación en la que entra en juego una codificación. Ineliminable en la estructura funcional del sistema, necesaria para la digitalización, la cuantización trae consigo este ruido que es bastante diferente al ruido aditivo que se presenta en el canal de transmisión.

Para medir este defecto definiremos una razón señal a ruido de cuantización al que denotaremos por S/N_q . Varios métodos se han desarrollado para mejorar esta relación, aspectos técnicos que en otros tiempos no eran posibles habían detenido el desarrollo de este tópico de las comunicaciones, sin embargo el constante avance de la electrónica en general ha logrado hasta cierto grado maximizar la S/N_q mediante el uso de los modernos compansores, dispositivos que cuantizan no linealmente la señal analógica.

2.3.1. RELACIÓN SEÑAL A RUIDO DE CUANTIZACIÓN EN NIVELES IGUALMENTE ESPACIADOS

Téngase una señal aleatoria que tiene una curva de probabilidad dada separemos los niveles de cuantización una distancia común a (Fig.2.3.1.1).

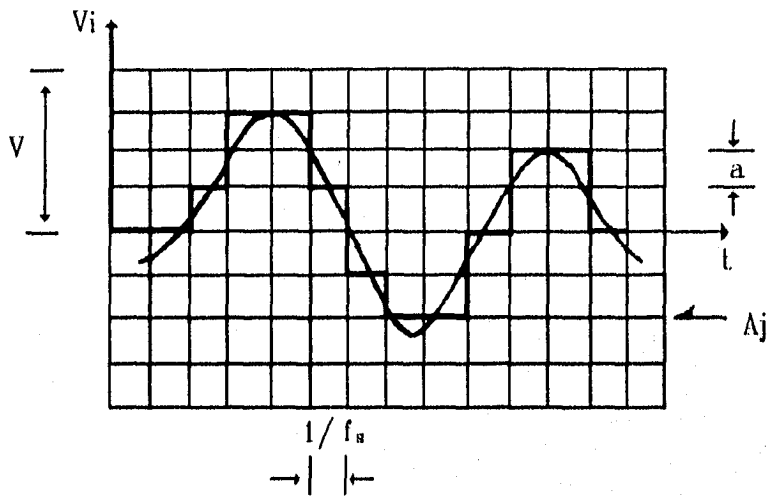


Fig. 2.3.1.1. Cuantización uniforme.

La diferencia de la señal cuantizada a la señal original la definiremos como el error que se cometa en este proceso; se calculará en base a considerar un nivel de decisión $\pm a/2$ (Fig.2.3.1.2) así suponiendo que todos los niveles continuos de la señal se repiten en la región de incertidumbre con la misma ocurrencia en un largo período de tiempo, el error cuadrático medio será:

$$E(\epsilon^2) = \frac{1}{a} \int_{-a/2}^{a/2} \epsilon^2 d\epsilon$$

$$E(\epsilon^2) = \frac{a^2}{12} \quad (2.3.1.1)$$

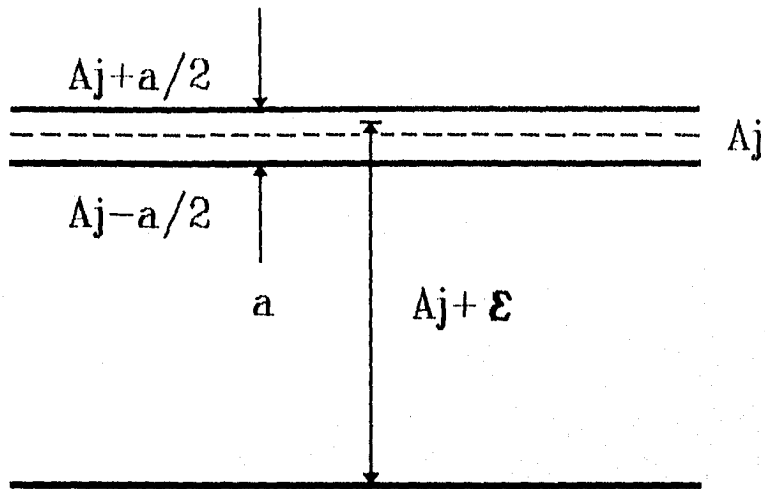


Fig. 2.3.1.2. Región de incertidumbre para el cuantizador.

Hagamos corresponder a la variancia de la distribución probabilística σ^2 la potencia media de la señal que aparece a la entrada del cuantizador y V la amplitud de la señal pico. Por lo que :

$$\frac{S}{N_q} = \frac{\sigma^2}{E(\epsilon^2)} \quad (2.3.1.2)$$

y

$$a = \frac{2V}{N} \quad (2.3.1.3)$$

Sustituyendo esta última ecuación en la ec. 2.3.1.1

$$E(\epsilon^2) = \frac{V^2}{3N^2} \quad (2.3.1.4)$$

la que absorbida por la ec. 2.3.1.2 resulta en

$$\frac{S}{N_q} = 3N^2 \frac{\sigma^2}{V^2} \quad (2.3.1.5)$$

o

$$\frac{S}{N_q} = 3N^2 \sigma_x^2 \quad (2.3.1.6)$$

σ_x = potencia media de entrada normalizada

Debido a que la S/N_q es proporcional a σ^2 en la ec. 2.3.1.5, es indicativo que la potencia de la señal es más evidente en cuanto se aumenta la intensidad de ésta, ya que la potencia media del ruido permanece como una constante.

2.3.2. RELACIÓN SEÑAL A RUIDO DE CUANTIZACIÓN EN SEÑALES CON COMPANSIÓN

Para eliminar la dependencia de la S/N_q a la intensidad y obtener con ello un valor fijo para un amplio rango dinámico la utilización de la estadística es útil. La existencia de niveles de amplitud que ocurren con más frecuencia que otras admite la posibilidad de cuantizar no linealmente, es decir los niveles de separación variarlos conforme a la probabilidad de ocurrencia de amplitud. Esto en telefonía es muy usado ya que la voz es de naturaleza aleatoria y las señales más débiles ocurren más frecuentemente; es aquí donde la utilización de compansores entra en juego. La compansión es el método bajo el cual se comprime en amplitud los niveles superiores y se expanden los niveles inferiores de una señal analógica antes de cuantizarla linealmente, siguiéndose un procedimiento inverso en el receptor para recuperar la forma real de la información (Fig. 2.3.2.1). Por lo que, la compansión se utiliza como acrónico del proceso de compresión y expansión.

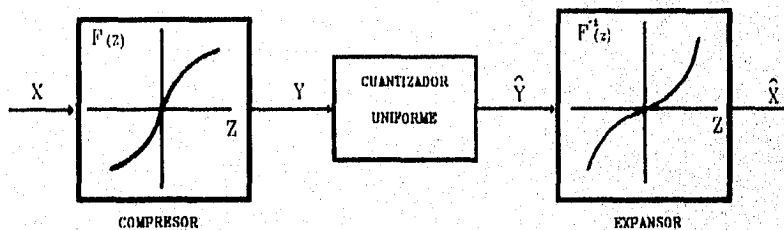


Fig. 2.3.2.1. El proceso de compansión

Actualmente existen dos técnicas de uso mundial para realizar la compresión conocidas como ley A y ley μ dadas por las siguientes ecuaciones :

$$F(x) = \frac{\ln(1 + \mu X)}{\ln(1 + \mu)} \quad 0 \leq X \leq 1 \quad (2.3.2.1)$$

$$\frac{1 \cdot \ln A x}{1 \cdot \ln A} \quad \frac{1}{A} \leq X \leq 1$$

(2.3.2.2)

$$\frac{Ax}{1 \cdot \ln A} \quad 0 \leq X \leq \frac{1}{A}$$

$X = \frac{X'}{V}$ = amplitud de la señal normalizada.

La primera aplicada en Estados Unidos, Canadá y Japón, la segunda en Europa. Ambas con simetría impar respecto al punto $X=0$ con parámetros típicos $\mu = 255$ y $A = 100$ para cuantización a 8 bits, $\mu = 100$ y $A = 87.6$ para 7 bits. La Fig.2.3.2.2 muestra las curvas de compresión para la ec. 2.3.2.1. con distintos valores de μ para $X > 0$.

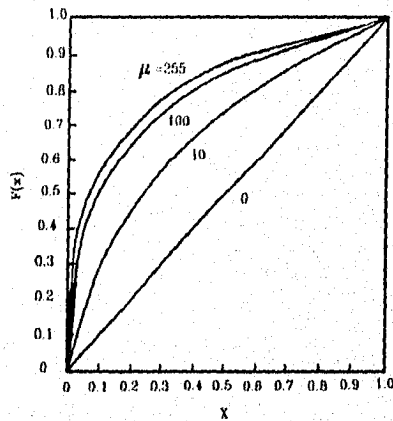


Fig. 2.3.2.2. Ley μ de compresión.

Un análisis para la compansión con ley μ demuestra que la S/N_q está dado por:

$$\frac{S}{N_q} = \frac{3N^2}{\ln(1 + \mu)^2} \frac{1}{1.2E \frac{(|X|)}{\mu\sigma_x^2} \cdot \frac{1}{\mu^2\sigma_x^2}} \quad (2.3.2.3)$$

$E(|X|)$ y σ_x son parámetros estadísticos de la señal que encaja en alguna función de densidad probabilística (T.2.3.2.1).

Función	LAPLACIANA	GAUSSIANA	UNIFORME
$f(x)$	$\frac{1}{\sqrt{2}\sigma_x} e^{-\frac{\sqrt{2} x }{\sigma_x}}$	$\frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{x^2}{2\sigma_x^2}}$	$\frac{1}{2A} - A \leq X \leq A$
$\frac{2 E(X)}{\sigma_x}$	1.414	1.6	1.732

T.2.3.2.1. Funciones de densidad de probabilidad.

Para la señal específica de voz, la función de densidad de probabilidad inherente a ella es la Laplaciana; es interesante comparar los resultados obtenidos con la cuantización lineal y no lineal. La Fig.2.3.2.3 gráfica las ecs. 2.3.1.6 y 2.3.2.3 para una señal analógica en f.d.p. Laplaciana, con $\mu = 255$ y 128 niveles de cuantización.

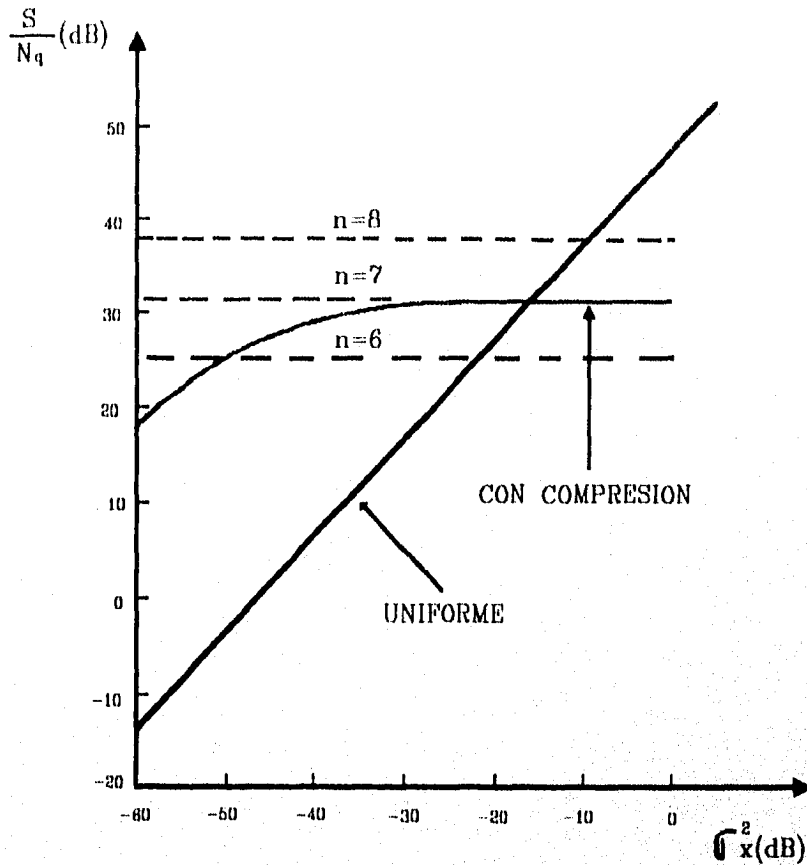


Fig. 2.3.2.3. Relación señal a ruido para señales con cuantización uniforme y con compresión.

El amplio rango dinámico de trabajo logrado con la compansión al mantenerse la S/N_q casi constante sobre 40 dB según la Fig.2.3.2.3, degrada hasta cierto punto el proceso de espaciamento uniforme cuya curva ilustra la susceptibilidad de la S/N_q a cambios de amplitud de la señal analógica.

En la práctica la relación señal a ruido para ley μ se calcula por :

$$\frac{S}{N_q} (dB) = 6n - 8.5 \quad \mu = 100 \quad (2.3.2.4.a)$$

$$\frac{S}{N_q} (dB) = 6n - 10.1 \quad \mu = 255 \quad (2.3.2.5.b)$$

Esto se basa en considerar constantes las curvas de la S/N_q con compansión para todo el rango dinámico. De esta manera es evidente el incremento de la relación señal a ruido conforme el aumento del número de bits deseados para la cuantización. En la Fig.2.3.2.3 se ilustra lo anterior por medio de las curvas punteadas para $\mu = 255$.

Finalmente obsérvese el mejoramiento en la calidad de la cuantización cuando se usa un compresor en una señal de prueba senoidal (Fig.2.3.2.4).

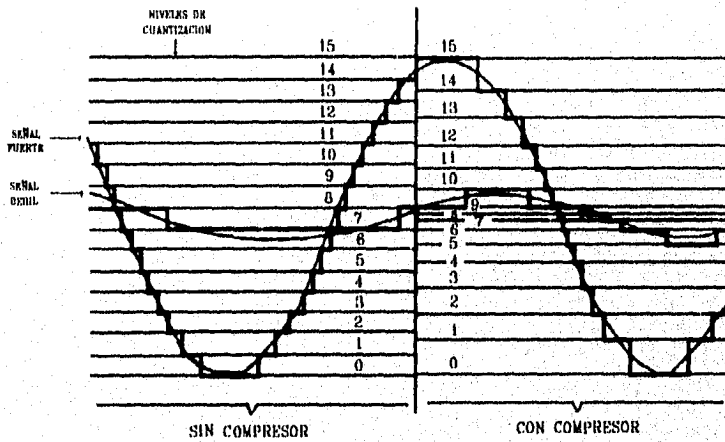


Fig. 2.3.2.4. Señal senoidal cuantizada con compresor y sin compresor.

2.4. EL SISTEMA BELL SYSTEM T1 CARRIER

En esta sección es conveniente analizar un ejemplo vivo de aplicación de la teoría del PCM a la transmisión de voz. El sistema T1 de la Bell System aún cuando viejo a tiempo de funcionamiento se refiere, presenta los principios fundamentales en cuestión de multiplexaje y transmisión de canales de voz vía repetidores en línea.

2.4.1. FUNCIONAMIENTO

Este sistema multiplexa 24 canales de voz con 7 bits para codificación de una muestra. El sistema es diseñado para transmitir frecuencias de voz hasta 4 KHz y por tanto 8000 muestras por segundo son requeridas. Un cuadro representa el total de muestras tomadas en todos los canales, cada cual es representada por 8 bits, según se muestra en la Fig. 2.4.1.1. El bit 8 es usado para supervisión y señalización. Existe un total de 193 bits en cada cuadro, y así la línea del sistema T1 opera a $193 \times 8000 = 1544000$ bits/seg.

El último bit de cuadro, el 193, es usado para establecer y mantener sincronización. La secuencia de esos bits 193 para distintos cuadros es establecido por la lógica de la terminal receptora. Si esta secuencia no sigue tal código patrón, entonces la terminal detecta que la sincronización ha sido perdida.

Esto es, si existe un deslíz en la sincronización, los bits examinados podrían ser de hecho bits de los canales de voz, los cuales deben presentar baja probabilidad de formar el código. Se ha encontrado que una secuencia teórica de 0 1 0 1 0 1..... es adecuada y esto ha funcionado en la práctica. Al circuito de cuadro le toma 0.4 a 0.6 mseg. detectar la asincronía, además de un tiempo de recuadro con valor de 50 mseg. en el peor de los casos.

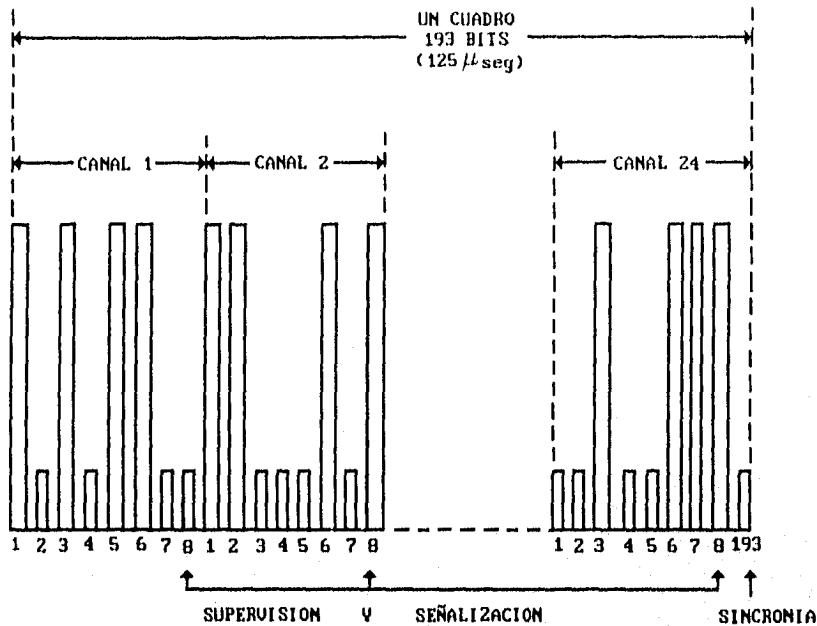


Fig. 2.4.1.1. Distribución de los canales de voz en el sistema Bell System T1 carrier.

En el caso T1 CARRIER la sincronización es realmente simple dado que los errores de codificación que puedan surgir se hacen transparentes cuando aparecen ante el oído humano. En otros sistemas más sofisticados en los que se requiere la transmisión de bloques de datos, es necesario el uso de códigos detectores de errores, estos debidos a la influencia del ruido a través de la línea.

Los niveles de señal para este esquema T1 PCM no son igualmente espaciados, sino que a amplitudes bajas la nivelación es más angosta que para las amplitudes mayores. Esto produce mejor reproducción del bajo volumen de voz.

2.4.2. REPETIDORES REGENERATIVOS

La principal razón por la cual existe la alta velocidad de bits que puede ser lograda en circuitos con pares de cable usando PCM es que los repetidores son puestos a intervalos frecuentes para reconstruir la señal.

En la mayoría de los sistemas PCM trabajando en la actualidad los repetidores son colocados a intervalos entre 1 y 5 Km. El sistema T1 usa repetidores a intervalos de 1.8 Km, el cual es el espacio entre bobinas de carga empleadas para transmisión analógica: los repetidores reemplazan a las bobinas de carga. Estos repetidores reconstruyen 1544000 pulsos por segundo.

Un repetidor regenerativo realiza tres funciones básicas: remodelado, temporización y regeneración. Cuando un pulso arriba al repetidor, este llega atenuado y distorsionado, por lo cual debe pasar por un preamplificador y ecualizador con el fin de remodelarlo para el proceso de detección. Un filtro remueve la componente de DC. Un circuito recuperador de tiempo provee una señal para muestrear el pulso al punto óptimo para decidir si es un bit con valor 1 ó 0. Este circuito de tiempo controla la regeneración del pulso saliente y asegura que es enviado al tiempo y ancho del pulso correcto.

Como es mostrado en la Fig.2.4.2.1 los pulsos transmitidos ocupan la mitad de un time slot. Un pulso positivo o negativo representa un 1 y la ausencia de pulso denota un 0. Esto concentra la energía de la señal alrededor de 772 Khz en vez de 1.544 MHz cuando una cadena de 1's es transmitida.

El sistema T1 utiliza una señal de 3 niveles (A, 0, -A) conocida como señal bipolar RZ, con el fin de provocar un margen contra variaciones antes de la detección y decisión. Transmisiones digitales más avanzadas utilizan señales polares, un pulso positivo denota un 1 y un pulso negativo un 0. Una señal polar es más eficiente; mayor velocidad en bits puede ser transmitida sobre un canal dado.

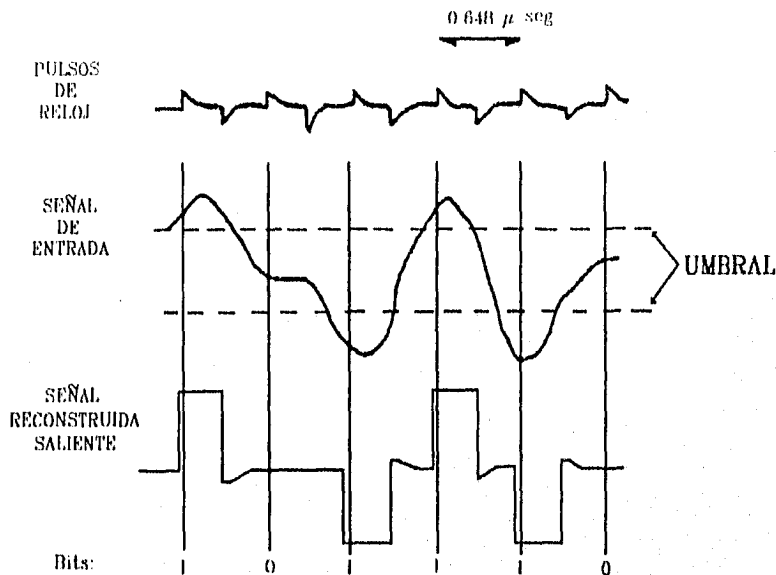


Fig. 2.4.2.1. Reconstrucción de una señal digital en un repetidor del sistema T1.

2.5. PROCESO DE COMPRESIÓN Y EXPANSIÓN (COMPANDING) DE UNA SEÑAL

Si nos concentramos en una transmisión telefónica, podremos tener un orador débil (bajo volumen), donde el cuantizador deberá interpretar los sonidos del nivel más bajo, es decir, debe tener un grano muy fino para conservar la inteligibilidad. Granos más ordinarios pueden ser soportados por los picos de voz del mismo orador o por otro más fuerte (alto volumen).

Los intervalos son clasificados en tamaño, fino a amplitudes pequeñas (al centro del rango), y progresivamente más ordinarios con el incremento en amplitud (hacia los extremos del rango).

Los intervalos son clasificados en tamaño, fino a amplitudes pequeñas (al centro del rango), y progresivamente más ordinarios con el incremento en amplitud (hacia los extremos del rango).

Esto puede ser realizado indiferentemente, por cuantización no uniforme desempeñada directamente en el codificador, o como en la práctica, con el uso de niveles de cuantización cónica (tapered), donde las señales son primeramente comprimidas no linealmente en amplitud para forzar a todas las señales a estar situadas dentro de un rango especificado. Esta característica de compresión es típicamente de forma logarítmica. Los niveles de cuantización uniforme son entonces aplicados para esta señal comprimida, el efecto es para proporcionar más niveles de cuantización a los niveles de señal más pequeña, como si el espaciamiento de niveles de cuantización hayan sido reducidos a niveles de señal más bajos. En el decodificador la señal es expandida a la amplitud que tenía con una compensación a través de una forma logarítmica inversa.

La combinación del proceso de compresión y expansión es denominado companding. Para la voz, los efectos del companding o niveles de cuantización cónica no uniforme es proporcionar una parte de señal a ruido cuantizado sobre un rango dinámico mucho más ancho de señales. Algunos tipos de cuantización no uniforme pueden también ser derivados por traslación digital después de la codificación de grano fino uniforme. En la Fig. 2.5.1. se muestran tres métodos en forma de diagrama a bloques para obtener cuantización uniforme.

2.5.1. CUANTIZACIÓN LOGARÍTMICA O LEYES COMPANDING

La ley companding se realiza cuando introducimos los niveles de cuantización cónica, y como ya se explicó, en la transmisión de voz de diferentes oradores usando las mismas condiciones, el rango de intensidad de voz variará hasta 40 dB. Una de las razones para la utilización del espaciamiento no uniforme de los niveles de

cuantización, es debido a que con cuantización uniforme no resulta fácil cubrir efectivamente el rango dinámico producido por los diversos oradores con diferentes tonos suaves y tono fuertes. La otra razón es que, la distribución de amplitud de señal es raramente uniforme. En voz, las pequeñas amplitudes son mucho más probables que las grandes.

Con cuantización no uniforme en las Ecs.(2.6.1.) a (2.6.3.), la probabilidad P_i es muy diversa en valor y las contribuciones del ruido de los diversos intervalos difieren. Por lo que será conveniente que el ruido total de la Ec.(2.6.3.) sea reducido, para ésto se incrementa el tamaño de los intervalos que son improbables de ser ocupados y reduciendo el tamaño de aquellos intervalos cuyos usos son altamente probables.

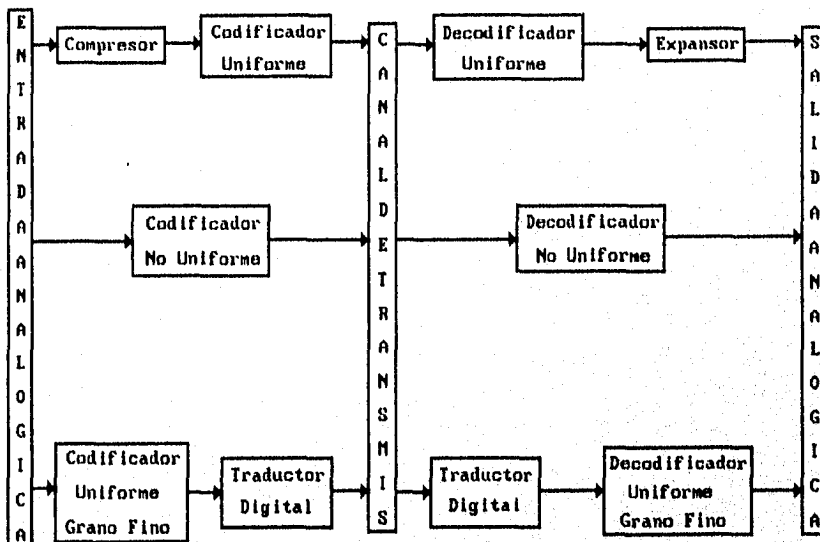


Fig. 2.5.1. Métodos de obtención de cuantización no uniforme.

La no uniformidad de la distribución de amplitud tiene un efecto adicional. Una señal codificada transporta la máxima información solamente si todos los niveles son equiprobables.

Así, la dispersión en P_i reduce el flujo de información y por tanto parece benéfico el uso de grandes intervalos en esas partes del rango donde la densidad de probabilidad es menor. La máxima información no coincide en general con el mínimo ruido.

En la práctica, una señal tiene una distribución de amplitud alineada y normalmente estará sujeta a una mezcla de cuantización uniforme y logarítmica. La potencia del ruido no puede ser menor que el debido a cada tipo de cuantización, pero tampoco puede exceder a la suma de los dos. Podemos hablar de cuatro leyes companding que son :

- i) Ley A o ley logarítmica con segmentos tangenciales lineales.
- ii) Ley cuasi - logarítmica o ley μ .
- iii) Ley seno hiperbólico inverso.
- iv) Ley hiperbólica.

Para la cuantización de las señales de voz con calidad aproximadamente uniforme sobre un amplio rango de volumen se puede usar alternativamente las leyes companding. Cualquier ley companding debe recurrir a una relación lineal o cercana a lo lineal a niveles bajos.

- i) La ley A es definida por la siguiente ecuación :

$$y = \frac{Ax}{(1 + \log A)} \quad 0 \leq x \leq \frac{1}{A}$$

$$y = \frac{1 + \log Ax}{1 + \log A} \quad \frac{1}{A} \leq x \leq 1 \quad (2.5.1.1)$$

La pendiente inversa de la curva de compresión, la cual define los intervalos del quantum es :

$$\frac{dx}{dy} = \frac{K_1}{A} \quad 0 \leq x \leq \frac{1}{A}$$

$$\frac{dx}{dy} = K_1 x \quad \frac{1}{A} \leq x \leq 1 \quad (2.5.1.2)$$

donde $K_1 = 1 + \log A$.

Para una cuantización uniforme esta pendiente debe ser unitaria. Así, el cuanta sobre el segmento lineal central, el cual tiene un efecto dominante sobre señales pequeñas es disminuído por el factor para esta ley A/K_1 , que es la ventaja companding. Este factor tiene un valor típico de 25 dB. Otra medida útil de cuantización logarítmica es el tamaño del cuanta, medida sobre una escala logarítmica (el logaritmo de la relación entre dos umbrales próximos).

Para calcular y evaluar el ruido cuantizado para cualquier señal deseada y ley companding, podemos partir de la Ec.(2.6.3.) de donde:

$$\frac{1}{2} N \delta_i = \left(\frac{dx}{dy} \right)_{x=x_i}$$

es la pendiente aproximada de una tangente de la curva companding y el error queda:

$$E^2 = \frac{1}{12} \sum P_i \delta_i^2 = \frac{1}{3N^2} \sum P_i \left(\frac{dx}{dy} \right)_{x=x_i}^2 \quad (2.5.1.3)$$

para muchos niveles, la sumatoria se puede aproximar por una integral, con lo que la ecuación anterior será :

$$E^2 = \frac{1}{3N^2} \int_{-1}^1 P(x) \left(\frac{dx}{dy} \right)^2 dx \quad (2.5.1.4)$$

El SNR es por tanto :

$$SNR = \frac{1}{3N^2} \frac{\int P(x) \left(\frac{dx}{dy} \right)^2 dx}{\int P(x) x^2 dx} \quad (2.5.1.5)$$

el ruido debido a cuantización logarítmica ideal será :

$$SNR = \frac{K^2}{3N^2} \quad (2.5.1.6)$$

Para el caso de la ley A sustituimos en la (2.5.1.2) y nos resulta un error cuadrático:

$$E^2 = \frac{K_1}{3N^2} \left[\int_{-1}^1 P(x) x^2 dx + \int_{-\frac{1}{A}}^{\frac{1}{A}} P(x) (A^{-2} - x^2) dx \right] \quad (2.5.1.7)$$

donde la primera integral es la potencia de la señal en términos de la distribución de probabilidad (S^2). La relación señal a ruido se obtiene dividiendo la Ec.(2.5.1.7) entre S^2 :

$$\frac{E^2}{S^2} = \frac{K_1^2}{3N^2} \left[1 + \frac{1}{S^2} \int_{-\frac{1}{A}}^{\frac{1}{A}} P(x) (A^{-2} - x^2) dx \right] \quad (2.5.1.8)$$

donde el primer término es conocido como el ruido debido a una cuantización logarítmica ideal y el segundo término es debido a la ley logarítmica A.

ii) La ley μ es definida por la ecuación :

$$y = \frac{\log(1 + \mu x)}{\log(1 + \mu)} \quad (2.5.1.9)$$

lo cual no es estrictamente lineal ni estrictamente logarítmica, pero será aproximadamente lineal a niveles bajos ($X < \mu^{-1}$) y aproximadamente logarítmica a niveles altos ($X > \mu^{-1}$). El tamaño de la etapa logarítmica no es constante, pero a niveles altos tiende asintóticamente a :

$$K_2 = \log(1 + \mu) \quad (2.5.1.10)$$

Haciendo sustituciones como en el caso anterior la Ec.(2.5.1.4) en términos de la ley μ el error medio cuadrático será :

$$E^2 = \frac{K_2^2}{3N^2} \int P(x)(x + \mu^{-1})^2 dx \quad (2.5.1.11)$$

y la relación señal a ruido es :

$$\frac{E^2}{S^2} = \frac{K_2^2}{3N^2} \left[1 + \frac{1}{\mu^2 S^2} + \frac{2}{\mu S^2} \int P(x) dx \right] \quad (2.5.1.12)$$

iii) La ley logarítmica SENO HIPERBOLICO INVERSO se define por:

$$y = \frac{\sinh^{-1} Cx}{\sinh^{-1} C} \quad 0 \leq x \leq 1 \quad (2.5.1.13)$$

Donde el tamaño de la etapa logarítmica a niveles altos tiende asintóticamente a :

$$K_3 = \sinh^{-1} C \quad (2.5.1.14)$$

Con las sustituciones apropiadas en la Ec.(2.1.5.4.) en términos de la ley seno h^{-1} el error medio cuadrático será :

$$E^2 = \frac{K_3^2}{3N^2} \int P(x)(x^2 + C^2) dx \quad (2.5.1.15)$$

en donde la relación señal a ruido es :

$$\frac{E^2}{S^2} = \frac{K_3^2}{3N^2} \left[1 + \frac{1}{C^2 S^2} \right] \quad (2.5.1.16)$$

iv) La ley COMPANDING HIPERBÓLICA es diferente de las tres anteriores porque no hace una aproximación asintótica, sin embargo, el quantum se incrementa uniformemente con el nivel de la señal, y puede ser formulada completamente en un codificador secuencial con una entrada:

$$x = \frac{y}{(m+1) - my}$$

por lo que la ley hiperbólica es definida por :

$$y = \frac{(1+m)x}{1+mx} \quad (2.5.1.17)$$

Donde el tamaño de la etapa logarítmica no tiene una asíntota horizontal como en las otras leyes, pero tiene un valor mínimo de :

$$K_4 = \frac{4m}{1+m} \quad (2.5.1.18)$$

El ruido cuantizado es :

$$E^2 = \frac{K_4^2}{3N^2} \int P(x) \left[\frac{(1 - mx)^2}{4m} \right]^2 dx \quad (2.5.1.19)$$

y la relación señal a ruido será :

$$\frac{E^2}{S^2} = \frac{K_4^2}{3N^2} \left[\frac{(1 - g)^2}{4ms} \right]^2 \quad (2.5.1.20)$$

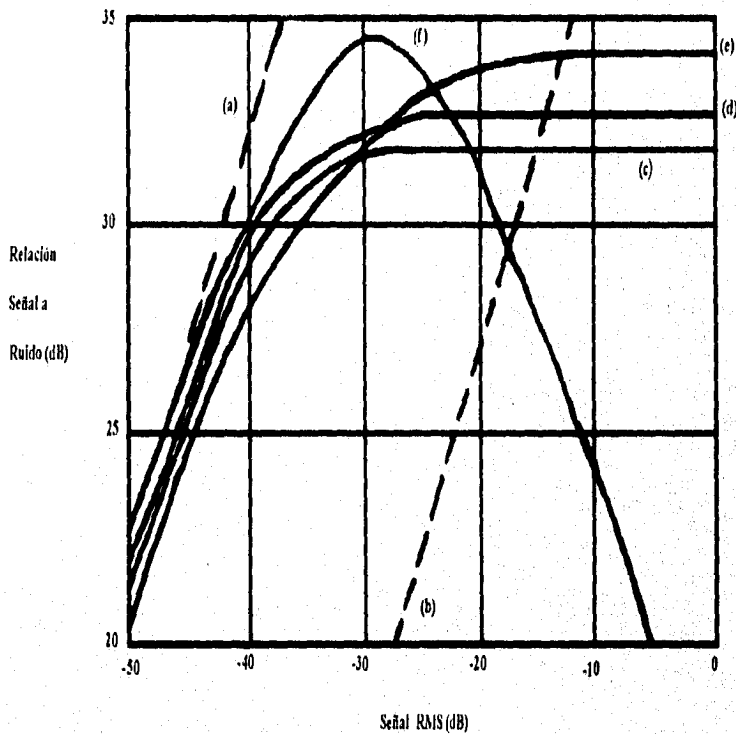


Fig. 2.5.2. Relación señal a ruido para la distribución exponencial (como voz) con 128 niveles de cuantización para las cuatro leyes companding.

En la tabla 2.5.1.1. se hace una comparación tangible y práctica de las propiedades de cada una de las leyes companding. En la Fig.(2.5.2.) se muestra la relación señal a ruido calculada para la distribución exponencial (como voz) con 25 dB de ventaja companding y 128 niveles de cuantización o 7 dígitos binarios.

La línea punteada del lado derecho (b), muestra la proporción señal a ruido debido a cuantización uniforme con el mismo número de niveles, esta línea tiene pendiente unitaria porque el ruido es constante. La línea punteada del lado izquierdo (a) es el valor asintótico más bajo en todas las leyes companding. La diferencia entre a y b es la ventaja companding. La curva (c) define la ley A; la curva (d) define la ley seno hiperbólico inverso; la curva (e) define a la ley μ ; y la curva (f) define la ley hiperbólica. Dichas curvas exhiben un incremento progresivo con el incremento de la relación señal a ruido y las tres primeras (c, d, e) se mantienen aproximadamente constantes a partir del valor máximo, mientras que la cuarta cae.

La conclusión de lo anterior es que cualquiera de las leyes logarítmicas permiten una extensión del rango de voz. Además existen diversos métodos de incorporación de cuantización no uniforme en el diseño de un codificador y todas estas leyes companding pueden ser realizadas.

TIPO DE LEY	ECUACION DE DEFINICION	$\frac{1}{x} \frac{dx}{dy} = \frac{1}{2}$ veces el tamaño de la etapa logarítmica		VENTAJA COMPANDING	
		EXPRESION GENERAL	VALOR MINIMO O ASINTOTICO		
Ley A o Ley Logarítmica con segmentos tangenciales Lineales	$y = \frac{Ax}{1 + \log A} \quad 0 \leq x \leq \frac{1}{A}$	$\frac{1 + \log A}{Ax} \quad 0 \leq x \leq \frac{1}{A}$	$K_1 = 1 + \log A$ $Ax \geq 1$	$\frac{A}{1 + \log A}$	$\frac{A}{K_1}$
	$y = \frac{1 + \log Ax}{1 + \log A} \quad \frac{1}{A} \leq x \leq 1$	$1 + \log A \quad \frac{1}{A} \leq x \leq 1$			
Ley μ o Ley Cuasi - Logarítmica	$y = \frac{\log(1 + \mu x)}{\log(1 + \mu)}$	$\frac{(1 + \mu x) \log(1 + \mu)}{\mu x}$	$K_2 = \log(1 + \mu)$ $\mu x \rightarrow \infty$	$\frac{\mu}{\log(1 + \mu)}$	$\frac{\mu}{K_2}$
Ley Seno Hiperbólico Inverso	$y = \frac{\sinh^{-1} Cx}{\sinh^{-1} C}$	$\frac{(\sinh^{-1} C) \sqrt{1 + C^2 x^2}}{Cx}$	$K_3 = \sinh^{-1} C$ $Cx \rightarrow \infty$	$\frac{C}{\sinh^{-1} C}$	$\frac{C}{K_3}$
Ley Hiperbólica	$y = \frac{x + mx}{1 + mx}$	$\frac{(1 + mx)^2}{(1 + m)x}$	$K_4 = \frac{4m}{1 + m}$ $mx = 1$	$1 + m$	

Tabla 2.5.1.1 Propiedades de las leyes de Companding

2.6. EFECTO DEL RUIDO EN EL CANAL DE TRANSMISIÓN

Durante la transmisión de información por un canal dado, existen fuentes aleatorias de ruido que tienden a producir perturbaciones indeseables en nuestra señal de información. Así, definimos al ruido como una señal indeseable sin relación alguna con la señal deseada. Para un estudio analítico de los efectos del ruido y métodos de solución a este, existen dos figuras de mérito importantes: la probabilidad de error y la capacidad del canal. La primera es un indicador del grado de inseguridad en la señal recibida. La segunda, se refiere a las limitaciones físicas del canal, debidas al ancho de banda y la potencia del ruido, las cuales permiten una velocidad máxima de transmisión para una cierta rapidez de errores.

2.6.1. CARACTERÍSTICAS DEL RUIDO

De las varias fuentes de ruido que existen, haremos la siguiente clasificación : a) ruido producido por el hombre, b) perturbaciones naturales y c) ruido de fluctuación. El ruido producido por el hombre es aquel que se debe a fenómenos bien conocidos de nuestra vida diaria y que provienen de fuentes tales como contactos defectuosos, artefactos eléctricos, radiación por ignición, alumbrado fluorescente, etc. Tal tipo de ruido puede evitarse eliminando la fuente que lo produce.

El ruido natural errático, el cual es producido por fenómenos de la naturaleza misma, puede proceder de relámpagos, tormentas eléctricas en la atmósfera, ruido intergaláctico, etc.

El ruido de fluctuación aparece en los sistemas físicos de transmisión tales como elementos resistivos, dispositivos semiconductores, válvulas de vacío, etc. Básicamente existen dos tipos de ruido de fluctuación : el ruido de disparo y el ruido térmico. El primero debido a la emisión aleatoria, recombinación y difusión de portadores en semiconductores. El segundo se debe al movimiento aleatorio de los electrones libres en medios conductores tales como resistores, y que de una manera u otra es dependiente de la temperatura.

En cuestión del ruido, es necesario englobar de cierta manera todas las fuentes y todos los efectos de ellas en un modelo matemático único que sirva para iniciar un estudio analítico del problema de la transmisión de errores. El modelo de ruido que se utiliza es el ruido blanco gaussiano, el cual presenta una densidad espectral de potencia constante para todas las bandas de frecuencia y una amplitud que varía de acuerdo a una distribución de probabilidad gaussiana con valor esperado igual a cero. Esto es, la amplitud del ruido estará supeditada a la siguiente expresión:

$$f(n) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{n^2}{2\sigma^2}} \quad (2.6.1.1)$$

donde $n(t)$ es la amplitud del ruido aleatorio y σ^2 es la varianza del ruido la cual asumimos conocida. La Fig.2.6.1.1 muestra un oscilograma típico de ruido en voltaje y su función de densidad probabilística del tipo gaussiano.

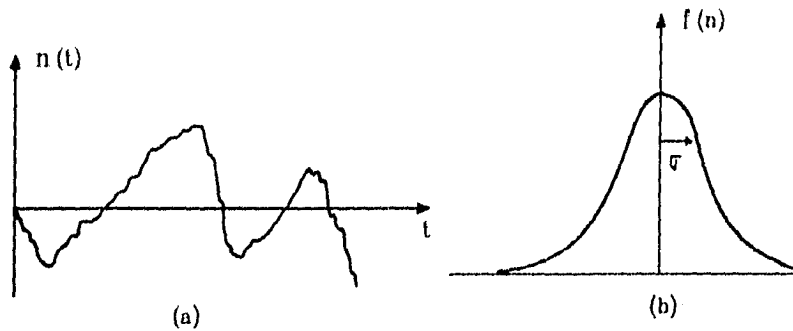


Fig. 2.6.1.1. Ruido blanco gaussiano. a) oscilograma, b) función de densidad de probabilidad.

2.6.2 PROBABILIDAD DE ERROR

Analicemos por separado los tipos de transmisión digital en banda base y por onda portadora. Para las señales banda base consideremos el caso unipolar mostrado en la Fig.2.6.2.1 en la cual se hace referencia a un tren de pulsos cuadrados con ruido aditivo del canal superimpuesto. Esta señal de voltaje debe ser detectada por la circuiteria y dar una decisión del nivel enviado. El análisis teórico demuestra que la probabilidad de error para una señal unipolar está dada por la siguiente ecuación :

$$P_e = \frac{1}{2} \left(1 - \operatorname{erf} \frac{A}{2\sqrt{2}\sigma} \right) \quad (2.6.2.1)$$

$\frac{A}{\sigma}$ = relación amplitud de señal a ruido r.m.s.

función de error = erf x.

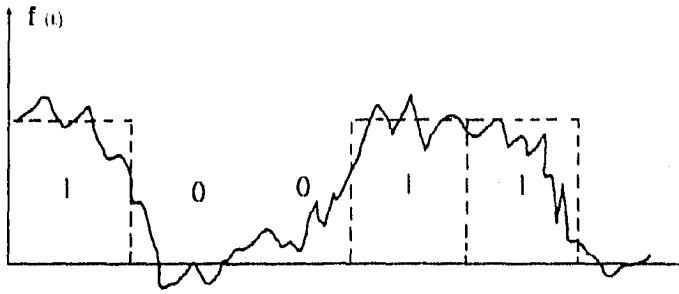


Fig. 2.6.2.1. Señal unipolar con ruido blanco gaussiano.

La función de error a la que hace referencia esta ecuación, se encuentra tabulada en manuales de matemáticas. La probabilidad de error se encuentra graficada para varios valores de A/σ en la Fig.2.6.2.2. Dado que es conveniente hablar en términos de relaciones de potencia, resultados que se basan en una ecuación semejante a la anterior, han llevado a obtener curvas como las que se indican en la Fig.2.6.2.3 para el caso de señales polares y bipolares también (pulso⁺ = 1, pulso⁻ = 0; pulso⁺ = 1, no pulso = 0, respectivamente).

En los códigos de línea asumidos hasta ahora, dos niveles de señalización (binarios) han sido usados. En aplicaciones donde el ancho de banda es limitado y rapidez de datos más altos son requeridos, el número de niveles puede ser incrementado manteniendo la misma rapidez de señalización. La rapidez de datos es obtenida de la siguiente manera :

$$R = \left(\frac{1}{T}\right) \log_2 L \quad (2.6.2.2)$$

donde:

L = número de niveles

T = intervalo de señalización

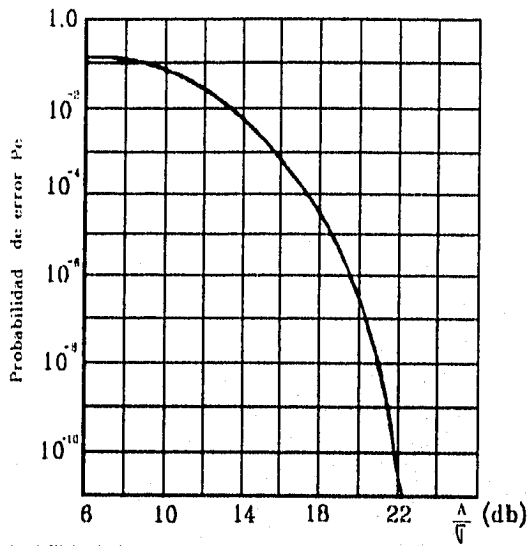


Fig. 2.6.2.2. Probabilidad de error para una señal unipolar.

Para este caso, la probabilidad de error está dada por :

$$P_e = \frac{1}{\log_2 L} \left(\frac{L-1}{L} \right) \operatorname{erfc}(z) \quad (2.6.2.3)$$

donde

$$z = \frac{(\log_2 L)^{\frac{1}{2}}}{L-1} \left[\frac{E_b}{N_0} \right]^{\frac{1}{2}}$$

$\frac{E_b}{N_0}$ = relación energía por bit a densidad de potencia del ruido.(especificada

sobre un lado de la banda espectral)

erfc $z = 1 - \operatorname{erf} z$ = función de error complementaria.

$$\frac{E_b}{N_o}$$

Está relacionada a la relación en potencia señal a ruido por:

$$\frac{S}{N} = 2 \frac{E_b}{N_o} \log_2 L \quad (2.6.2.4)$$

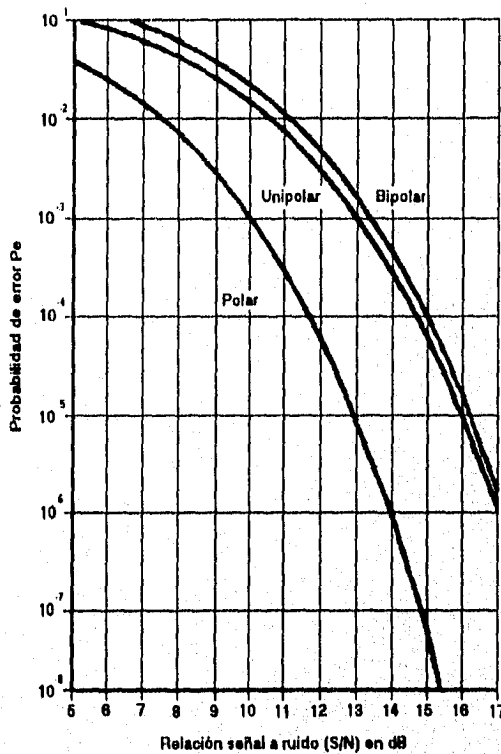


Fig. 2.6.2.3. Probabilidad de error para señales unipolar, polar y bipolar.

Las gráficas de P_e para señales multinivel se muestran en la Fig.2.6.2.4. Pasemos al caso de las señales moduladas en onda portadora. Para este tipo de señales, las gráficas de P_e en función de C/N para los modems más frecuentemente usados, son mostrados en la Fig.2.6.2.5. C/N representa la relación de la potencia media de la

portadora a la potencia media del ruido.

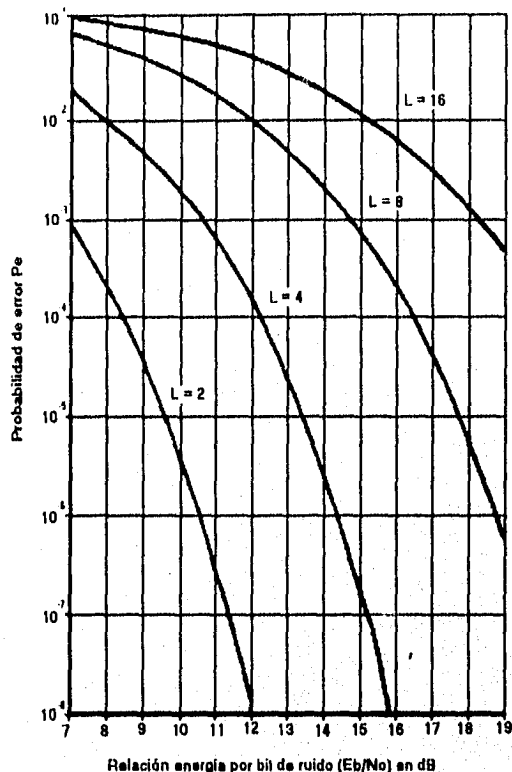


Fig. 2.6.2.4. Probabilidad de error para señales multinivel.

2.6.3. CAPACIDAD DE CANAL

Todos sistema o canal de transmisión es capaz de transmitir determinada cantidad de información por unidad de tiempo. Esto se conoce como capacidad del canal y se expresa en bits/seg. Las limitantes a la cantidad de información, son el ancho de banda permitido y la potencia de la señal de ruido. En 1949 Shannon desarrolló una fórmula que permite expresar la relación existente entre los factores anteriores considerando un canal de transmisión con ruido blanco gaussiano, la cual está dada por :

$$C = W \log_2 \left(1 + \frac{S}{N} \right) \quad (2.6.3.1)$$

donde :

C = capacidad del canal

W = ancho de banda

$$\frac{S}{N} = \text{relación en potencia señal a ruido}$$

Esta ecuación indica que es posible realizar un compromiso entre el ancho de banda y la potencia de la señal transmitida. Así, por un incremento lineal del primero debemos tener un decremento exponencial del segundo y viceversa. La Fig.2.6.3.1 muestra el comportamiento de la ec. en función de W. Obsérvese que la capacidad del canal se satura cuando el ancho de banda se incrementa fuertemente; $n_b/2$ (referida en 2.6.2.4.) es la densidad de potencia del ruido (especificada sobre los dos lados de la banda espectral). La ley de Shannon establece además que, si la rapidez de transmisión es menor que la capacidad del canal, la rapidez de error tiende a ser cero; mientras que si la transmisión es más veloz que la capacidad del canal, los errores se incrementan rápidamente.

Dado que Shannon desarrollo su ecuación sin considerar un sistema de comunicación específico, es conveniente conocer la ecuación particular que se ha desarrollado para PCM. Se ha encontrado que la capacidad del canal para este tipo de modulación está dada por :

$$C = W \log_2 \left(1 + \frac{12 S}{K^2 N} \right) \quad (2.6.3.2)$$

donde K representa el número de veces que la separación entre niveles supera al valor r.m.s. del ruido (A/σ en la Fig.2.6.2.2).

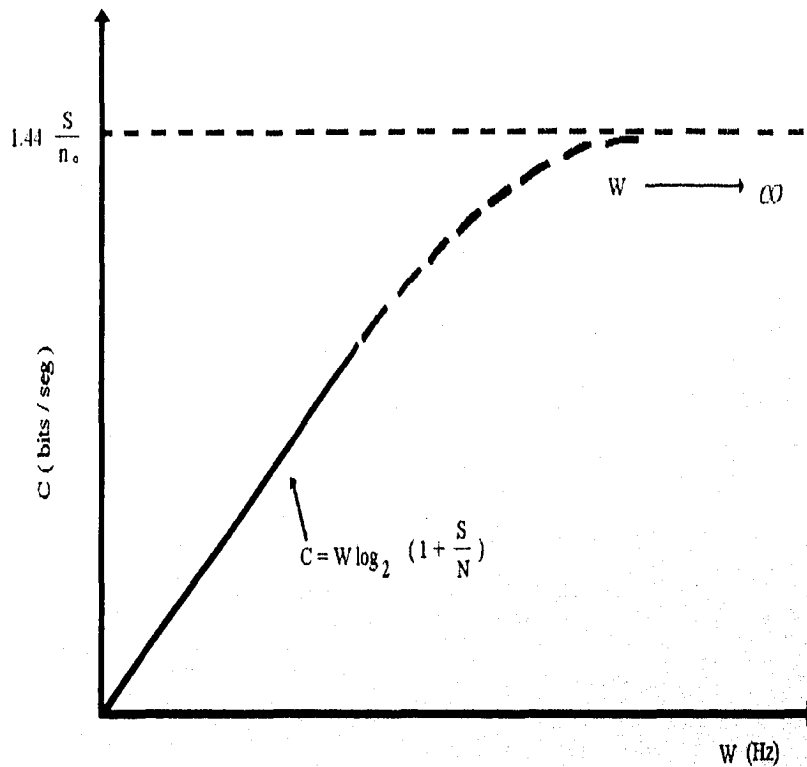


Fig. 2.6.3.1. Capacidad del canal de transmisión.

De esta ecuación se observa el factor $K^2/12$ e indica el número de veces en potencia que hay que aplicar a la señal enviada para que esta tenga la misma capacidad de canal expresada por la ley Shannon. Así, para una probabilidad de error de 10^{-6} , obtenemos refiriéndonos a la Fig.2.6.2.2. un valor de $K = 9.44$ por lo que se requiere 7.4 veces la potencia en la señal PCM que la requerida para un canal de Shannon.

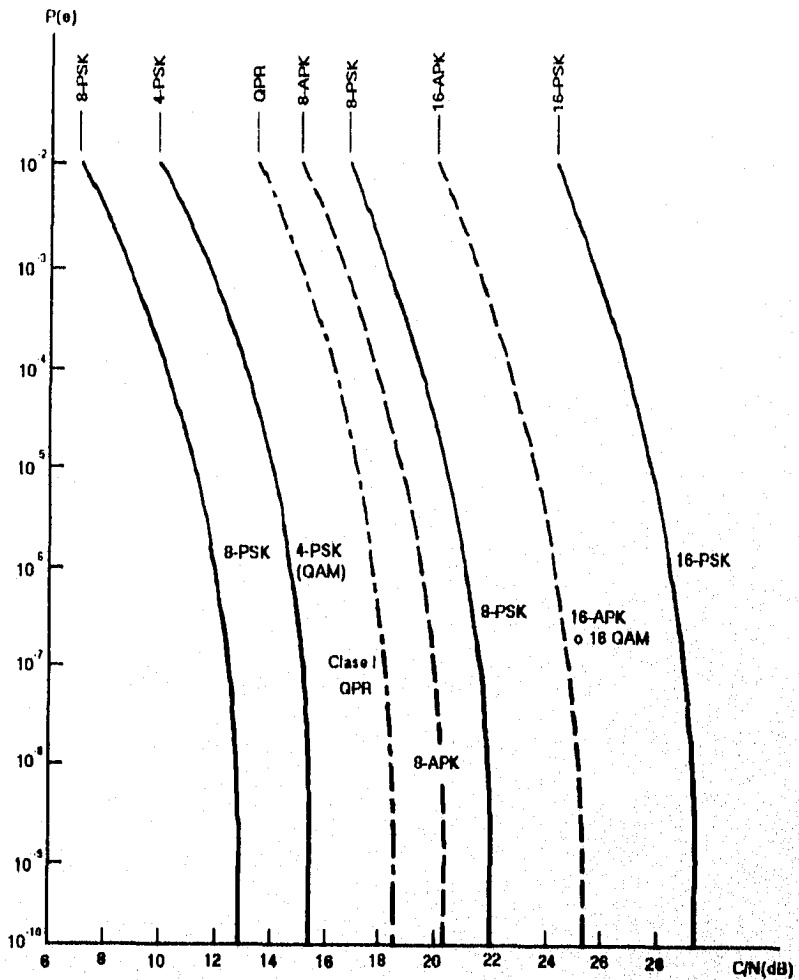


Fig. 2.6.3.2. Probabilidad de error en transmisión de señales por onda portadora.

2.7. FILTROS PASA - BAJAS

Estos tipos de filtros permiten el paso a los componentes espectrales de frecuencia de una señal que sean menores a una frecuencia de corte ω_c atenuando aquellas que se encuentren fuera de este rango. Este filtro tiene aplicación en el PCM tanto en la transmisión como en la recepción. Para el primer caso, se requiere para forzar a la señal de información a ser limitada en banda; en el segundo caso se utiliza para alisar la salida cuantizada del D/A.

Estos dispositivos pueden ser implementados a base de circuitos pasivos o activos de p - orden, donde el orden indica la capacidad de rapidez de atenuación para frecuencias fuera de la banda. Dadas las ventajas de los filtros activos sobre los pasivos, se analizará un tipo de estos, el VCVS.

2.7.1. FILTRO ACTIVO PB - VCVS

El filtro activo paso bajas VCVS (voltaje controlled - voltage source) utiliza una malla pasiva y una etapa amplificadora; se aplica retroalimentación entre las dos etapas (Fig.2.7.1.1.).

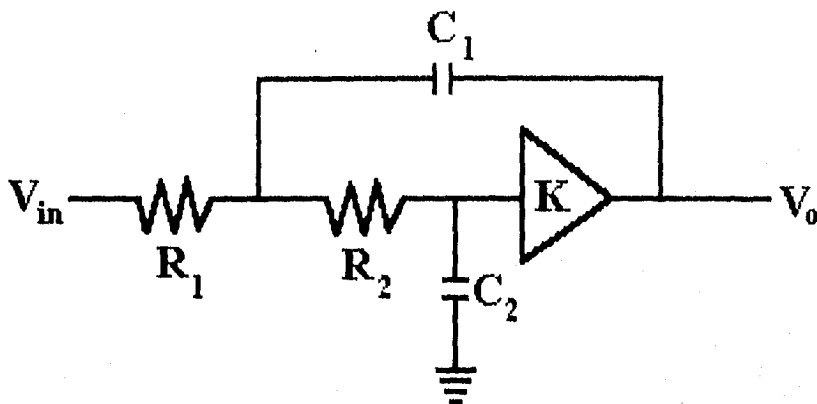


Fig. 2.6.1.1. Filtro activo paso - bajas VCVS.

Su función de transferencia en base a la variable compleja s , está dada por :

$$\frac{V_o(s)}{V_{in}(s)} = \frac{1}{s^2 + \left(\frac{1}{R_2 C_1} + \frac{1}{R_1 C_1} + \frac{1-K}{R_2 C_2} \right) s + \frac{1}{R_1 R_2 C_1 C_2}} \quad (2.7.1.1)$$

donde K = ganancia de la etapa amplificadora. La ec.(2.7.1.1) se puede indicar como:

$$\frac{V_o(s)}{V_{in}(s)} = \frac{\omega_o^2}{s^2 + 2\zeta\omega_o s + \omega_o^2} \quad (2.7.1.2)$$

donde :

$$\omega_o = \sqrt{\frac{K}{R_1 R_2 C_1 C_2}} \quad (2.7.1.2.a)$$

$$2\zeta\omega_o = \left(\frac{1}{R_2 C_1} + \frac{1}{R_1 C_1} + \frac{1-K}{R_2 C_2} \right) \quad (2.7.1.2.b)$$

La respuesta en frecuencia del filtro paso bajas se obtiene haciendo $s = j\omega$, de donde:

$$H(\omega) = \frac{1}{1 + 2\zeta \frac{\omega}{\omega_o} + \left(\frac{\omega}{\omega_o}\right)^2} \quad (2.7.1.3)$$

donde :

ω_o = frecuencia de ocurrencia del doble polo.

ζ = razón de amortiguamiento.

La Fig. (2.7.1.2) muestra la representación de Bode para la ec. (2.7.1.3) con los diagramas de magnitud y fase, para distintos valores de ζ . A partir del valor que tome la razón de amortiguamiento ζ , es posible obtener los filtros de 2^a orden Bessel, Butterworth y Chebyshev. La tabla (2.7.1.1) indica los valores de ζ para el diseño de estos filtros, así como el sobrenivel que se obtiene.

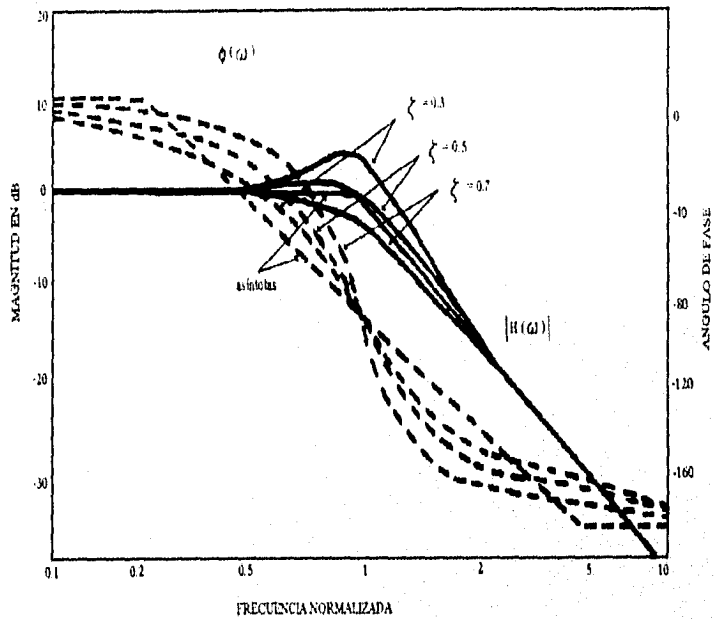


Fig. 2.6.1.2. Representación de Bode para el filtro activo PB.

FILTRO PB 2º ORDEN	ζ
Bessel	0.8659
Butterworth	0.7072
Chebyshev (0.1 dB pico)	0.6516
Chebyshev (0.25 dB pico)	0.6179
Chebyshev (0.5 dB pico)	0.5789
Chebyshev (1.0 dB pico)	0.5228
Chebyshev (2.0 dB pico)	0.4431
Chebyshev (3.0 dB pico)	0.3833

Tabla 2.7.1.1. Tipos de filtros en función de ζ .

ESTA FESIS NO DEBE
SALIR DE LA BIBLIOTECA

CAPITULO 3

TÉCNICAS DE MODULACIÓN DIGITAL

La separación física entre emisor y receptor implica la existencia de un medio de transmisión por medio del cual viaje la información. Este medio o canal de transmisión puede ser un par de hilos, cable coaxial, guía de onda, el espacio abierto, etc., en los cuales es necesario aprovechar eficazmente la disponibilidad frecuencial o de tiempo que se nos permita.

La modulación permite acoplar las características de la señal viajera al canal de transmisión. Siendo en esencia la modulación una técnica por la cual se procesa una señal variando sus características conforme a datos suministrados, obtiene dos objetivos fundamentalmente: 1) mayor eficiencia en la radiación de información debido al traslado en frecuencia de la señal modulante a la de portadora, 2) mejor aprovechamiento del espectro de transmisión por medio de una buena distribución de frecuencias portadoras. En particular, la transmisión digital vía portadora es muy semejante a la transmisión analógica, inclusive más simple en algunos casos.

Se ha desarrollado ampliamente la investigación de codificación de la forma de onda, enfocada en diversa áreas. Técnicas en el dominio del tiempo para razones de bit de 16 a 32 Kbps de mediana y baja complejidad han sido extensamente estudiadas y refinadas para un máximo funcionamiento y robustez. Aproximaciones para razones de bit de 16 Kbps y menores con alta complejidad no se encuentran del todo bien refinadas, sin embargo, las variaciones básicas de estas técnicas no están todavía exploradas. Igualmente, en el dominio de la frecuencia se ha desarrollado la investigación para mejoras principalmente en técnicas de codificación de transformación adaptiva y las técnicas relacionadas. La elección de la técnica a utilizarse dependerá de la calidad deseada, la razón de datos a manejar, la complejidad o costo requerido, pero sobre todo de la aplicación del sistema.

Básicamente existen tres técnicas de modulación digital; modulación en amplitud (AM), modulación en frecuencia (FM) y modulación en fase (PM). En años recientes sistemas híbridos (AM-PM o APK) han recibido gran atención debido a la economía del ancho de banda que se logra. Cada una de estas técnicas poseen un buen

número de variantes, lo más relevante de ellas será discutida aquí.

3.1. TÉCNICAS DE COMPRESIÓN DE SEÑALES DE ONDA

La representación de las señales en forma digital es en la actualidad de interés fundamental. Existen muchas maneras para representar las señales de voz en forma discreta. Como se vio en el capítulo anterior, estas formas de representar la voz se clasifican en tres y son:

- REPRESENTACIÓN DE FORMA DE ONDA.
- REPRESENTACIÓN DE LA FUENTE O VOCODER.
- REPRESENTACIÓN HÍBRIDA.

La representación de forma de onda como su nombre lo indica, implica preservar simplemente la "forma de onda" de la señal de voz analógica a través de un proceso de muestreo y cuantización. El diagrama a bloques de la Fig.3.1. describe en forma general la representación de la forma de onda. En ella, se ilustra la forma de onda de la voz, que puede ser idealizada como una función continua variable con el tiempo, la cual se muestrea generalmente a intervalos periódicos en el tiempo, para producir una secuencia de muestras $x(n)$. Para obtener una representación digital, es necesario cuantizar las muestras en un conjunto finito de valores apropiados.

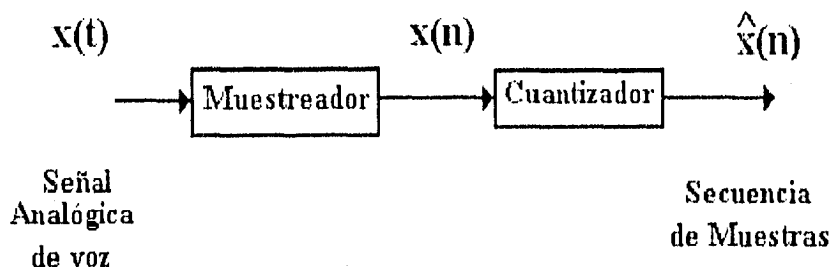


Fig. 3.1. Diagrama a bloques que describe la representación de la forma de onda.

3.2. MUESTREO Y CUANTIZACIÓN DE SEÑALES DE VOZ

En el capítulo anterior, se explicó el teorema de muestreo. Ahí se mostró que una señal analógica puede ser reconstruida en forma exacta partir de sus valores muestra, siempre y cuando la señal esté limitada en banda, y haya sido muestreada al rango de muestreo de Nyquist o mayor ($f \geq 2B$). Dado que nos interesa la representación digital de las señales de voz, es necesario considerar sus propiedades espectrales y además, tener en cuenta que las señales de voz no son inherentemente de banda limitada, aunque el espectro tiende a caer rápidamente a altas frecuencias.

En la práctica para asegurar la condición anterior, se introduce un filtro pasabajas de corte rápido antes del proceso de muestreo. Por otro lado, para representar exactamente todos los sonidos de voz, se requiere un rango de muestreo mayor de 20 KHz, debido a que el espectro de sonidos insonoros no caen apreciablemente sobre los 8 KHz. Sin embargo, en la mayoría de las aplicaciones este rango no se requiere (por ejemplo en telefonía), ya que dichas señales pueden ser especificadas con las tres primeras armónicas, por lo que solamente estaremos interesados sobre el rango del espectro en el límite de 3.5 KHz. Entonces, el rango de muestreo

práctico de las señales de voz de acuerdo al teorema de Nyquist, es de 8 KHz.

Para un fácil entendimiento es conveniente considerar en forma separada el proceso de muestreo y el de cuantización (en la práctica esto no es posible) en la representación digital de forma de onda. De esta manera se asume que la forma de onda de voz ha sido limitada en banda por un filtro pasabajas y después muestreada al rango apropiado dando una secuencia $x(n)$.

Esta señal $x(n)$ servirá entonces como una señal de entrada a cualquier proceso de PDS, o también para almacenarla directamente en una memoria digital, o bien para la transmisión sobre un canal de comunicaciones en forma digital. Los valores de las muestras son cuantizadas en un conjunto finito de amplitudes, ya que éstas pueden ser representadas por un conjunto finito de símbolos. Dando como resultado el proceso de cuantización y codificación. De la misma forma en que separamos el proceso de muestreo y el de cuantización, también es útil representar en forma separada el proceso de cuantización y codificación, esta acción se describe en la Fig.3.2. Aquí, la secuencia de muestras $x(n)$, son representadas por un conjunto finito de símbolos $c(n)$.

También en la figura podemos apreciar las dos etapas para realizar la representación: Una etapa de cuantización, la cual produce una secuencia de amplitudes cuantizadas $\hat{x}(n) = Q\{x(n)\}$, y una etapa de codificación, en la cual cada muestra cuantizada es representada por una palabra código $c(n)$. Todo lo anterior forma parte del codificador (Fig.3.2.a), donde Δ representa el tamaño de la etapa de cuantización en el cuantizador.

En el decodificador, la secuencia de palabras código $c'(n)$, se transforma dentro de una secuencia de muestras cuantizadas $\hat{x}'(n)$, que es descrito en la Fig.3.2.b. Si no hubo errores al recuperar $c'(n)$ entonces :

$$c'(n) = c(n)$$

por tanto

$$\hat{x}'(n) = \hat{x}(n).$$

En la mayoría de los casos, se usan números binarios para representar las muestras cuantizadas. Con palabras en código binario de m bits, es posible representar 2^m niveles diferentes de cuantización. La capacidad de información requerida para transmitir o almacenar la representación digital es :

$$I = (m)(f) = \text{RAZON DE BIT en bits/seg.} \quad (3.1.)$$

donde : f es el rango de muestreo (muestras/seg), y
 m es el número de bits/muestra.

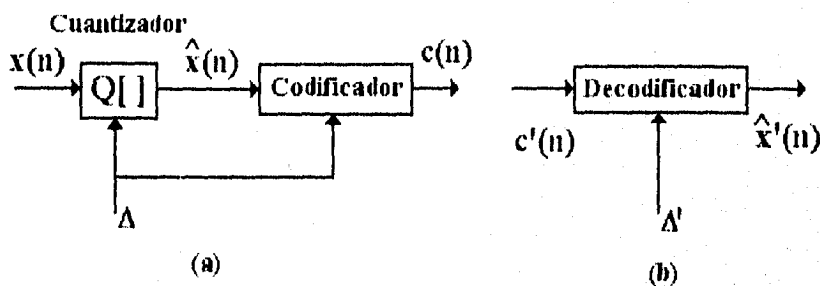


Fig. 3.2. Proceso de cuantización y codificación
a) Codificador b) Decodificador

En general, es deseable mantener la razón de bit tan baja como sea posible, a la vez que se mantiene un nivel deseable de calidad. Por lo que, para un ancho de banda de voz, el rango de muestreo está fijado por el teorema de muestreo, por tanto, el camino para reducir la razón de bit solamente es reducir el número de bits/muestra. Lo anterior se logra utilizando una gran variedad de técnicas, las que a continuación se explican.

3.3. MODULACIÓN POR CODIFICACIÓN DE PULSOS (PCM)

3.3.1. PCM LINEAL

Un método común y quizá la forma más simple de codificación de forma de onda es el PCM lineal, en el cual una señal analógica es uniformemente cuantizada, en este tipo los rangos y niveles de cuantización son distribuidos uniformemente. Existen dos clases de cuantizadores uniformes :

- CUANTIZADOR UNIFORME DE MEDIA SUBIDA "MID-RISER".
- CUANTIZADOR UNIFORME DE MEDIO PASO "MID-TREAD".

En el cuantizador uniforme de media subida, el origen aparece en el centro de una parte de la subida de la función como escalera (Fig.3.3.a.). Mientras que un cuantizador de medio paso, el origen aparece en el centro de una parte del escalón de la función como escalera (Fig.3.3.b.). Para un esquema de código binario, el cuantizador de media subida tiene el mismo número de niveles positivos y negativos, y está simétricamente posicionado alrededor del origen. El cuantizador de medio paso, tiene un nivel más negativo, en este caso uno de los niveles de cuantización es cero, mientras que el de media subida hay un nivel no cero. En la figura también, se muestran las palabras código asignadas que sirven para tener una ponderación numérica, interpretando las palabras código binario como una representación signo - magnitud. La ponderación es una palabra codificada en binario, en la cual nos da el signo (por ejemplo, "0" es un dígito que nos indica signo positivo, el dígito "1" nos indica signo negativo), y los dos bits restantes nos representan la magnitud entonces las muestras cuantizadas se relacionan

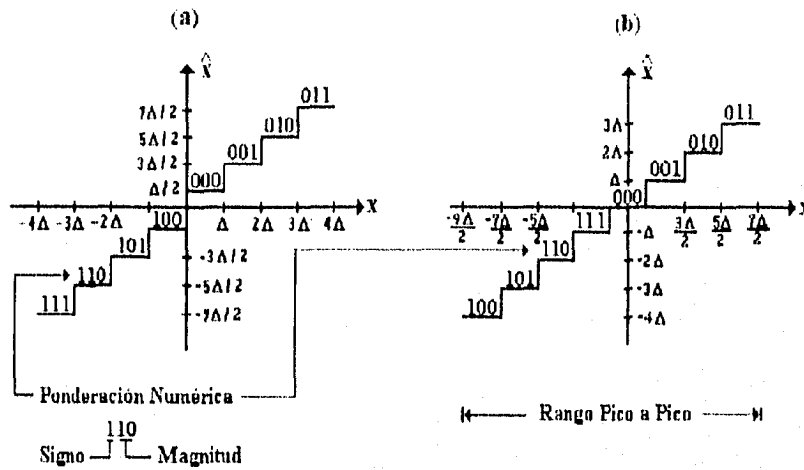


Fig.3.3. Representación de las características de los cuantizadores uniformes: a) Cuantizador de media subida; b) Cuantizador de medio paso.

con las palabras código por medio de lo siguiente:

$$\hat{x}(n) = \frac{\Delta}{2} \text{sign}c(n) + \Delta c(n) \quad (3.2.)$$

donde el sign { c (n) } es igual a +1 si el primer bit de c (n) es cero y -1 si el primer bit de c(n) es uno. En la Fig.3.3.b., las palabras codificadas en binario se interpretan como representaciones de dos complementos de 3 bits, donde las muestras cuantizadas son relacionadas con las palabras codificadas por la relación :

$$x(n) = \Delta \{ c(n) \} \quad (3.3.)$$

Para los cuantizadores uniformes, únicamente existen dos parámetros, el número de niveles (generalmente elegidos de la forma 2^m), y el tamaño de la etapa de cuantización Δ . Por lo que se elegirán adecuadamente Δ y m , para cubrir el rango de muestras de entrada. Por otro lado, la consideración que se hace para cualquier tipo de cuantización es la relación señal a ruido, de acuerdo a lo estipulado en el capítulo 2, el SNR para cuantización uniforme es :

$$\text{SNR}_{dB} = 10.8 + 6 m \quad (3.4.)$$

donde m es el número de dígitos binarios de las palabras de ponderación. Si en esta ecuación $m = 7$, tendremos un SNR de alrededor de 52.8 dB, en el cual es apropiado para proporcionar una calidad adecuada en un sistema de comunicaciones. Aunque para la presentación de la señal de voz como alta calidad, en un cuantizador uniforme se utilizará m igual a 11 o 12 bits, dado que no toma ventaja de las propiedades de la voz enunciadas en la tabla 3.3.1.; y en consecuencia no tiene capacidad de compresión de datos. En la cuantización uniforme, la relación señal a ruido es dependiente del nivel de la señal, por lo que será deseable que el error sea de varianza constante, es decir, independiente de la amplitud de la señal, obteniendo así un porcentaje de error constante que puede ser logrado con cuantización no uniforme o logarítmica.

A.- PRODUCCIÓN DE VOZ
1.- NO ESTACIONARIA EN TIEMPOS LARGOS
- Amplitud no estacionaria.
- Regiones [Sonoras/Insonoras/Silencios].
- Estructura del lenguaje y la fonética.
2.- ESTACIONARIA EN TIEMPOS CORTOS
- Estructura Formant (correlación en tiempos pequeños)
- Estructura de Tono "Pitch" (sonoros).
- Estructura de Ruidos (insonoros).
B.- PERCEPCIÓN DE VOZ
- Rango dinámico espectral local.
Enmascaramiento o disimulamiento auditivo.

Tabla 3.3.1. Propiedades Básicas de producción y percepción de voz.

3.3.2. PCM LOGARÍTMICO

Cuando se hace uso de una cuantización no uniforme de la amplitud de la señal, de acuerdo a una escala logarítmica en lugar de una escala lineal se habla de un PCM logarítmico, donde los niveles de cuantización están logarítmicamente espaciados. Como se puede apreciar en la Fig.3.4., el logaritmo de la entrada, es decir, $\log x(n)$ es cuantizada en lugar de la entrada $x(n)$ como en el caso lineal, es decir, realizándose el "companding", debido a que en el codificador las amplitudes de entrada son comprimidas por la función logarítmica antes de la cuantización, después son expandidas en el decodificador por la función exponencial. Esto conduce a una insensitividad a la amplitud de la señal, por tanto, el porcentaje de error es constante, y la relación señal a ruido es independiente de la varianza de la señal y depende solamente del tamaño de la etapa del cuantizador. De la figura se ve que :

$$y(n) = \ln\{ x(n) \} \quad (3.5)$$

por tanto :

$$x(n) = \exp\{ y(n) \} \text{sign}\{ x(n) \} \quad (3.6)$$

donde el $\text{sign}\{ x(n) \}$ es +1 si $x(n)$ es positivo y -1 si $x(n)$ es negativo. La magnitud logarítmica cuantizada es :

$$\begin{aligned} y(n) &= Q\{ \log | x(n) | \} \\ &= \log | x(n) | + \epsilon(n) \end{aligned} \quad (3.7)$$

pero $\epsilon(n)$ es independiente del $\log x(n)$, tenemos que:

$$\begin{aligned} \hat{x}(n) &= \exp\{ y(n) \} \text{sign}\{ x(n) \} \\ &= |x(n)| \text{sign}\{ x(n) \} \exp\{ \epsilon(n) \} \\ &= x(n) \exp\{ \epsilon(n) \} \end{aligned} \quad (3.8)$$

donde $\text{sign}\{ x(n) \} = \text{sign}\{ \hat{x}(n) \}$.

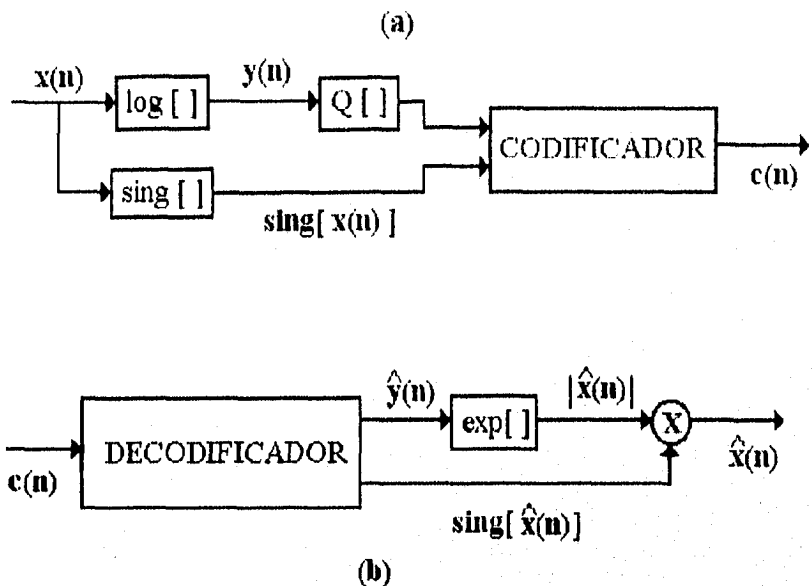


Fig. 3.4. Diagrama a bloques de la cuantización logarítmica a) Codificador b) Decodificador.

En la actualidad existen varias técnicas logarítmicas de cuantización, siendo las más utilizadas la ley μ y la ley A PCM. En dichas codificaciones se utilizan las características estáticas de la amplitud no estacionaria de la voz para conseguir buena calidad a razón de bit de 56 y 64 kbps.

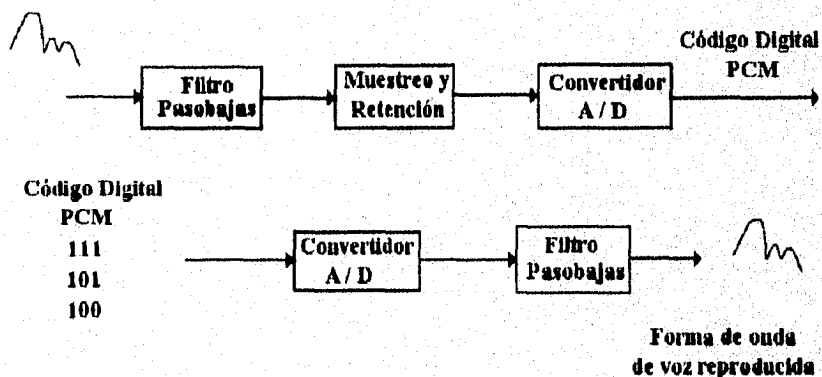


Fig. 3.5. Diagrama a bloques de un sistema de reproducción de voz que emplea PCM.

3.3.3. MODULACIÓN DELTA (DM)

Métodos opcionales para convertir las señales analógicas a un formato digital se han propuesto para algunas aplicaciones, con la esperanza de reducir el ancho de banda requerido, mejorar la eficiencia o reducir el costo. La modulación delta (DM) es uno de estos métodos y ha sido adoptado para algunas aplicaciones que requieren la transmisión de voz o de imágenes. En tales aplicaciones, existe un alto grado de redundancias en la información que se va a transmitir. La información pasada contiene también parte de información actual, como para poder predecirla aceptablemente bien, de manera que las señales nuevas sólo necesitan ser transmitidas si se producen cambios significativos en ella. Este proceso asume que las formas de onda de las señales de entrada serán bastante uniforme y con una pendiente predecible.

En la DM el rango de muestreo es muchas veces la razón de Nyquist para la señal de entrada, y esto se lleva a cabo para reducir el ruido que introduce el sistema y aparece a la salida del receptor. Como resultado del alto rango de muestreo, las muestras adyacentes están altamente correlacionadas. Este alto grado de correlación implica que seremos capaces de predecir mejor la muestra actual de las muestras pasadas, por tanto, la varianza de error de predicción será baja.

De este modo, la razón de bit de un sistema DM es simplemente igual al rango de muestreo. En la Fig.3.6., se muestra el sistema más simple de DM, que es la Modulación Delta Lineal (LDM). Donde el cuantizador tiene solamente dos niveles y el tamaño de la etapa está fija. Del diagrama a bloques, los valores de las muestras $x(n)$ se comparan con un valor muestra predicho denominado $\hat{x}(n)$, y la diferencia $d(n)$ se hace por el cuantizador. Obviamente, si $d(n)$ es pequeña durante la mayor parte del tiempo, de manera que la predicción es buena, se necesitarán pocos bits para representar esta señal diferencia.

El modulador delta, utiliza un cuantizador de dos niveles, de manera que sólo un bit se emplea para representar la señal. Los dos niveles del cuantizador son:

$c(n) = 0$ si el nivel de cuantización es positivo, y

$c(n) = 1$ si es negativo, entonces:

$$d(n) = \Delta \quad \text{si } c(n) = 0$$

$$= -\Delta \quad \text{si } c(n) = 1$$

En el receptor, la señal diferencia cuantizada se agrega a la salida del predictor para obtener una estimación discreta de $\hat{x}(n)$ de la señal muestreada deseada $x(n)$. En la DM el predictor empleado es generalmente una suma ponderada de un cierto número de muestras pasadas estimadas. Específicamente, siendo $x(n-1)$ la muestra previa, $\hat{x}(n-1)$ la estimación de ésta, $\hat{x}(n-2)$ la estimación de dos muestras anteriores etc., la forma general de un predictor lineal puede escribirse :

$$\hat{x}(n) = \sum_{l=1}^k h_l x(n-l) \quad (3.9.)$$

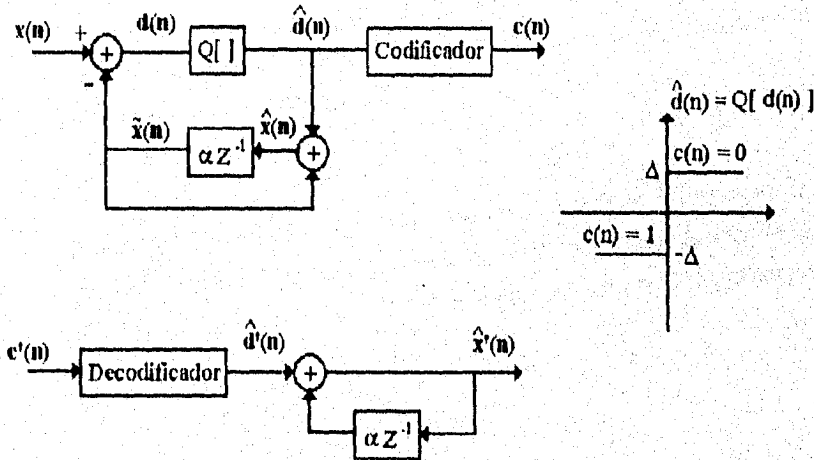


Fig. 3.6. Diagrama a bloques de un sistema de Modulación Delta; a) Codificador b) Decodificador

Los coeficientes h_l son los factores de ponderación. El predictor más simple es el que usa la estimación de la muestra previa sólo como una estimación de la muestra actual $x(n)$. Para éste caso se tiene :

$$x(n) = h_1 x(n-1) \quad (3.10.)$$

Los coeficientes h_l ($l=1,2,\dots,k$), pueden ser seleccionados para reducir en alguna medida los errores de modo que resulten tan pequeños como sea posible. Como en el caso de los sistemas PCM, el ruido también se introduce cuando se utiliza DM. A causa de que la señal diferencia es cuantizada en dos niveles, aparece un ruido granular o de cuantización similar al del PCM.

Cuando el rango de muestreo es demasiado bajo o el tamaño de la etapa elegida es también demasiado pequeño para perseguir una señal cambiando rápidamente, ocurre otro tipo de ruido llamado "ruido de sobrecarga"; éste ruido causa un error llamado "error de obediencia". El cual se adiciona al ruido de cuantización. Para apreciar la ocurrencia de estos dos tipos de ruido, en la Fig.3.7.a., observamos la operación de un modulador delta. Debido a que la cuantización es diferencial se usa la estimación de la muestras previas. En el tiempo $(n-1)$ de la figura el valor predicho $\hat{x}(n-1)$ es más pequeño que la señal $x(n-1)$. Debido a que :

$$d(n-1) = x(n-1) - \hat{x}(n-1)$$

es positivo, el tamaño de la etapa Δ es positiva. Esto es adicionado a $\hat{x}(n-1)$ para proporcionar la versión cuantizada de la entrada $\hat{x}(n-1)$. La señal predicha al tiempo de muestreo (n) es entonces :

$$\hat{x}(n) = \alpha_1 \hat{x}(n-1)$$

Si asumimos que α_1 es algo menor que uno, se obtiene $\hat{x}(n)$ como se ve en la figura. Si hacemos esto para cada punto, encontramos que la versión cuantizada $\hat{x}(n)$ tiende a seguir la curva $x(t)$, algunas veces situada encima y otras por debajo.

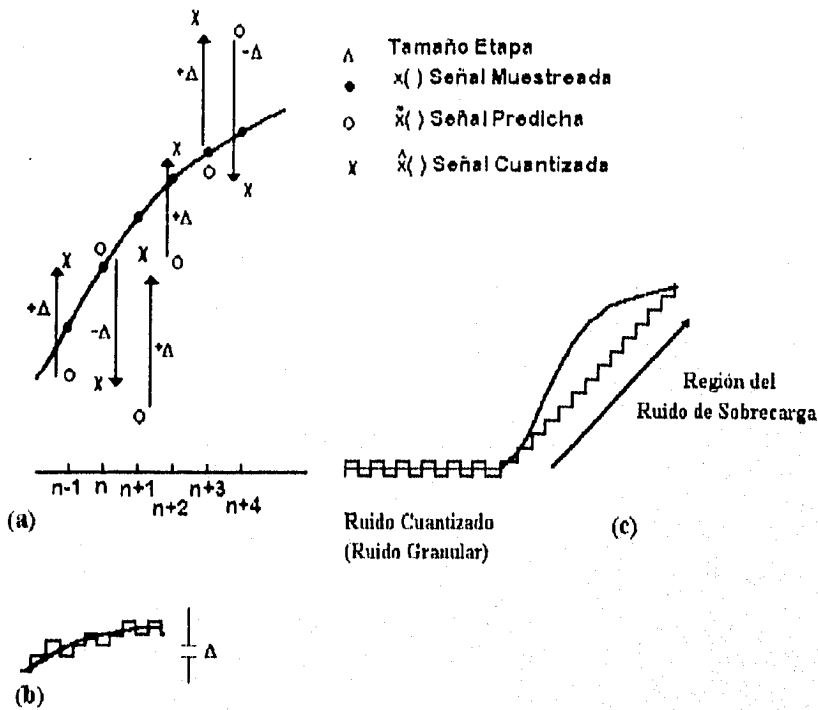


Fig. 3.7. Funcionamiento de un modulador delta :
 a) Forma gráfica de la operación del modulador delta.
 b) Forma de onda reproducida por un DM.
 c) Ruido de cuantización y ruido de sobrecarga.

El resultado es una señal recibida que persigue la señal original, pero introduce ruido cuantizado. La salida que aparecerá del modulador delta es como en la Fig.3.7.b., en el cual el ruido es proporcional al tamaño de la etapa Δ . Si se reduce Δ , el ruido también se reduce, pero si es demasiado pequeña, resulta el ruido de sobrecarga, y se debe a la inhabilidad del modulador para perseguir grandes cambios de señal.

La región de sobrecarga en la Fig.3.7.c., ocurre porque el tamaño de la etapa fija un límite máximo sobre la pendiente de la señal de entrada que el modulador puede seguir.

También Δ determina el error de pico, cuando la pendiente es muy pequeña. Por tanto, si incrementamos Δ , el ruido de sobrecarga disminuye, pero el ruido cuantizado aumenta, por lo que será deseable una elección óptima de Δ .

Abate, da la siguiente fórmula empírica para el tamaño de la etapa óptima :

$$\Delta_{opt} = \{ E [(x(n) - x(n-1))^2] \}^{1/2} \ln (2F_o) \quad (3.11.)$$

donde $F_o = F_s / 2B$, tal que F_s es el rango de muestreo del modulador y B la frecuencia de Nyquist. También da una relación señal a ruido en función del tamaño de la etapa variable normalizada:

$$SNR = \frac{\Delta}{\sqrt{E [x(n) - x(n-1)]^2}} \quad (3.12.)$$

Como se observa en la Fig.3.8., hay una localización del pico de la curva SNR para un valor de F_o . Si F_o se duplica, el SNR se incrementa a una razón de 9 dB. Al duplicar F_o duplicaremos la razón del bit debido a que :

$$RAZON DE BIT = F_s (1 \text{ bit}) = F_s = 2BF_o \quad (3.13.)$$

Comparándolo con la técnica PCM, tenemos que si duplicamos la razón de bit, duplicando el número de bits muestra, realizaremos un incremento de 6 dB por cada bit adicionado, de esta forma, el SNR incrementado con la razón de bit es mucho más dramático para el PCM que para el DM. Dado que solamente se utiliza un código de un bit en este sistema, se puede implementar con circuitos integrados simples analógicos y digitales, por lo que no requiere de sincronización de patrones de bit entre el transmisor y el receptor por tanto, la simplicidad de este sistema es su principal ventaja.

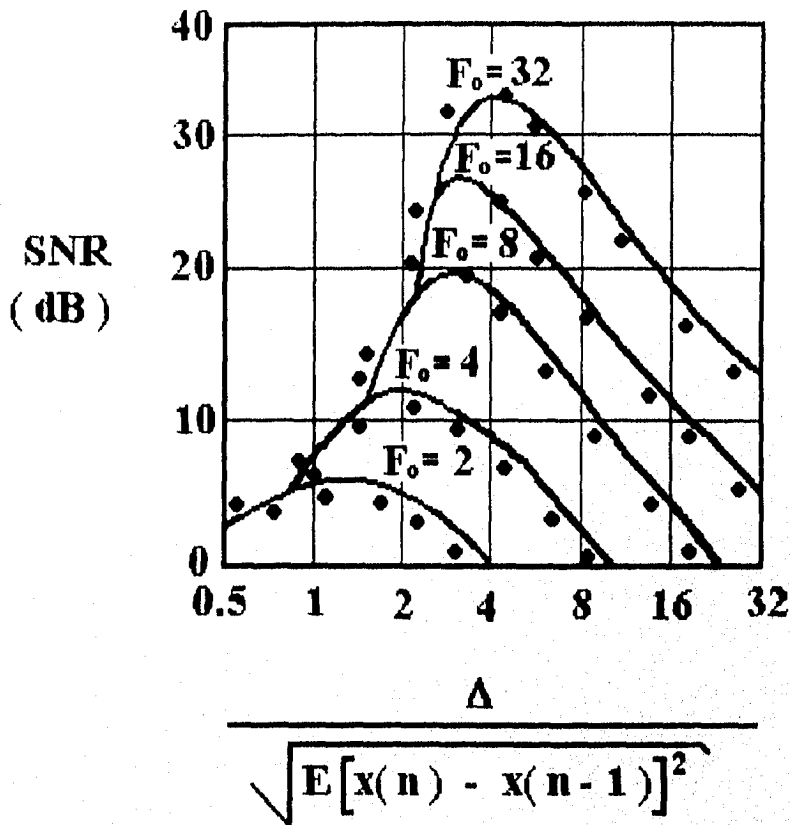
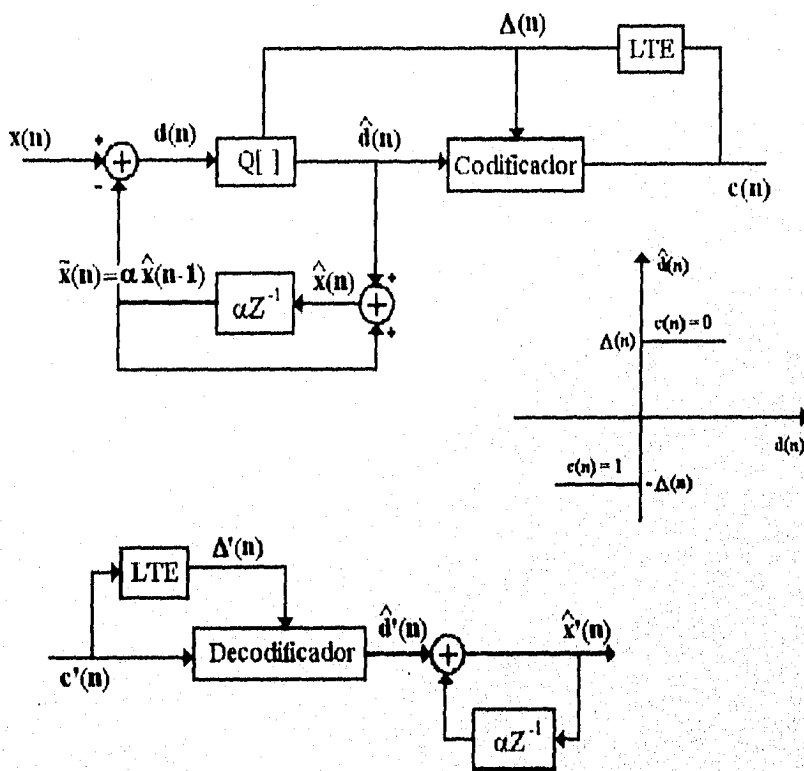


Fig. 3.8. Gráfica del SNR para DM en función del tamaño de la etapa variable normalizada.

Las limitaciones sobre el funcionamiento del sistema LDM, proviene principalmente de su cuantización tosca de la diferencia de la señal. Por lo que la cuantización adaptiva mejora grandemente el funcionamiento del modulador delta.

3.4. MODULACIÓN DELTA ADAPTIVA (ADM)

Una forma general de un sistema de Modulación Delta Adaptiva (ADM), es como el mostrado en la figura. Como en la Fig.3.9., la mayoría de estos esquemas son del tipo de cuantización adaptiva regenerativa o de realimentación, en el cual el tamaño de la etapa para un cuantizador de dos niveles es adaptado de la salida de la palabra código. En este sistema, se mantiene la ventaja de que no se requiere de la sincronización de patrones de bit.



(LTE) Lógica del Tamaño de la Etapa

Fig. 3.9. Diagrama a bloques de la Modulación Delta Adaptativa: a) Codificador b) Decodificador.

En el algoritmo ADM, el tamaño de la etapa obedece a la regla :

$$\Delta (n) = M \Delta (n-1) \quad (3.14.a)$$

$$\Delta_{\min} \leq \Delta (n) \leq \Delta_{\max} \quad (3.14.b)$$

En el cual el multiplicador es una función de las palabras código presentes y previas, $c (n)$ y $c (n-1)$. Esto es posible debido a que $c (n)$ depende solamente del signo de $d (n)$, el cual está dado por :

$$d (n) = x (n) - \alpha \hat{x} (n - 1) \quad (3.15.)$$

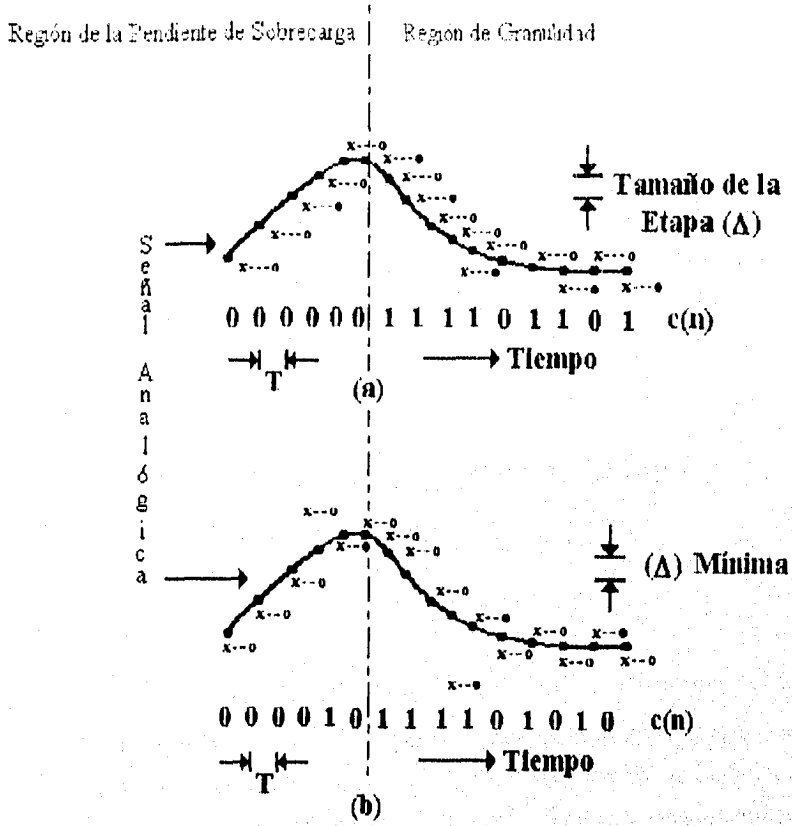
De este modo el signo de $d (n)$ puede ser determinado antes del proceso de la determinación del valor cuantizado actual $\hat{d} (n)$, el cual debe esperar la determinación de $\Delta (n)$ de la ec.3.14. El criterio para elegir el multiplicador del tamaño de la etapa de la ec. 3.1.4. es :

$$\begin{aligned} M = P > 1 & \quad \text{si} \quad c (n) = c (n-1) \\ M = Q < 1 & \quad \text{si} \quad c (n) \neq c (n-1) \end{aligned} \quad (3.16.)$$

En la Fig.3.10.a, se muestra una señal analógica procesada por medio de LDM, y en la Fig.3.10.b., la misma señal pero procesada ahora por ADM, donde :

- $x (n)$ es la señal de entrada.
- $\hat{x}(n)$ es la señal predicha.
- x $\hat{x}(n)$ es la señal cuantizada.

Con esta figura podemos apreciar con gran facilidad el resultado de ambos procedimientos, y la diferencia que provocan en la señal original. La región de la izquierda es la de sobrecarga (que es más pronunciada en LDM que en ADM), y a la derecha la región de granulidad. De las ecuaciones vemos que los parámetros a controlar del sistema ADM son: P , Q , Δ_{\min} , Δ_{\max} .



- $X(n)$ Señal muestreada
- $\tilde{X}(n)$ Señal predicha
- $\hat{X}(n)$ Señal cuantizada

Fig. 3.10. Señal analógica procesada mediante Modulación Delta :
 a) Tamaño de la etapa fija (LDM).
 b) Tamaño de la etapa adaptativa (ADM).

El tamaño mínimo de la etapa puede mantenerse tanto como la pendiente de la forma de onda se mantenga, o bien hacerla más pequeña que la requerida para el funcionamiento óptimo de un LDM, con lo que el ruido granular puede ser

grandemente reducido. Igualmente el tamaño de la etapa máxima puede ser hecha más larga que la pendiente máxima de la señal de entrada para reducir la pendiente del ruido de sobrecarga.

Los límites del tamaño de la etapa, serán elegidos para proporcionar el rango dinámico deseado para la señal de entrada. La relación $\Delta_{max} / \Delta_{min}$ (que controla el rango dinámico del cuantizador), será lo bastante grande para mantener un SNR alto sobre un rango deseado de los niveles de la señal de entrada. El tamaño mínimo de la etapa será tan pequeño como sea práctico para minimizar el ruido que se produce en la condición de canal desconectado.

3.5. COMPARACIÓN DE LOS SISTEMAS PCM LOGARÍTMICO, LDM Y ADM

En la Fig.3.11., se comparan sistemas LDM, PCM log ($\mu = 100$) y ADM ($P=1.5$), con un rango de muestreo de Nyquist ($2B=6.6$ Khz). En la figura se muestra que el sistema ADM es superior al LDM por 8 dB a 20 Kbps y la ventaja de SNR aumenta de 16 dB a 60 Kbps. Para el LDM, se observa que duplicando el rango de muestreo y la razón de bit, incrementa 6 dB mientras el ADM su correspondiente incremento es de 10 dB.

Comparando el sistema ADM y el PCM logarítmico, notamos que el ADM excede el funcionamiento del PCM log para una razón de bit por debajo de los 40 Kbps. Para razones de bit más alto el PCM log tiene un SNR más alto, pues como se aprecia, un ADM requiere 60 Kbps para llevar a cabo la misma calidad que con un PCM log de 7 bits que la realiza con una razón de bit de alrededor de 46 Kbps.

Para muchas aplicaciones ADM puede ser preferida al PCM log, inclusive a expensas de una razón de información ligeramente más alta. Debido a que el ADM mantiene la simplicidad del LDM, ya que la calidad mejorada del ADM se realiza solamente con un ligero incremento en complejidad, debido a que la adaptación del tamaño de la etapa se realiza usando la salida del bit corriente o actual.

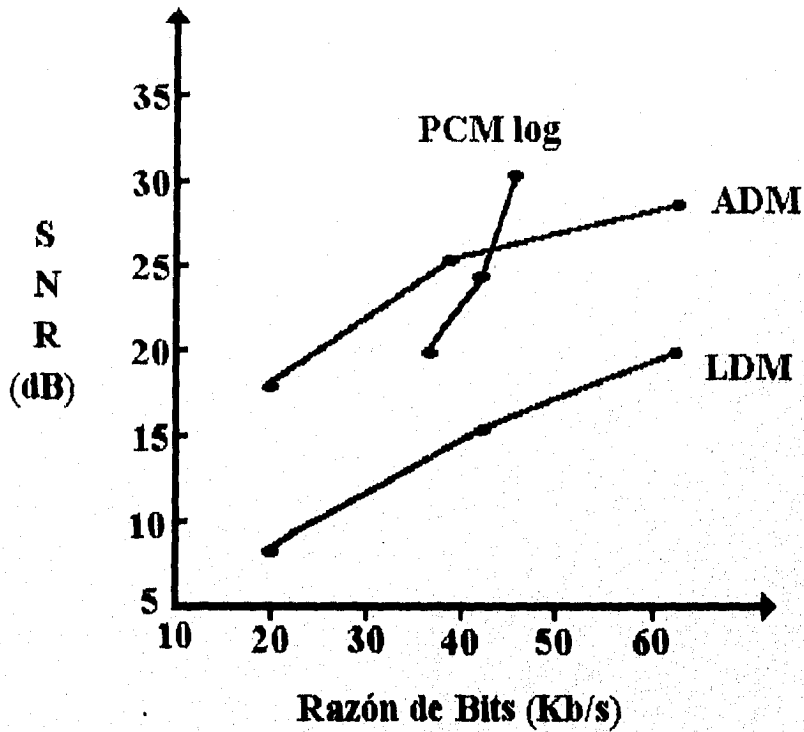


Fig. 3.11. Comparación gráfica SNR (dB) contra razón de bit (kbps) para los sistemas PCM logarítmico, LDM y ADM, con un rango de muestreo de 6.6 KHz.

3.6. MODULACIÓN DELTA CON PENDIENTE QUE VARÍA CONTINUAMENTE (CVSD)

La modulación delta con pendiente que varia continuamente (CVSD), es otro tipo de cuantización adaptiva en DM y éste también se basa en un esquema como el mostrado en la Fig.3.12., pero el tamaño de la etapa está definido por las ecuaciones:

$$\Delta(n) = \beta \Delta(n-1) + D_2$$

$$\text{si } c(n) = c(n-1) = c(n-2)$$

$$\Delta(n) = \beta \Delta(n-1) + D_1$$

de cualquier otra forma donde : $0 < \beta < 1$ y $D_2 > D_1 > 0$.

El tamaño de la etapa máximo se obtiene si la entrada al filtro del tamaño de la etapa es constante a D_2 como ocurre en un período prolongado de la pendiente de sobrecarga. El tamaño de la etapa mínimo, se obtendrá si el arreglo $c(n) = C(n-1) = C(n-2)$, no ocurre para períodos prolongados como en la condición de canal desconectado. El principio básico de funcionamiento del CVSD es como el del ADM, donde se incrementa el tamaño de la etapa en respuesta al patrón del bit actual, que indica la pendiente de sobrecarga. En el sistema CVSD una carrera de tres unos consecutivos o de tres ceros consecutivos causan un incremento de D_2 para ser adicionado al tamaño de la etapa. En ausencia de estas condiciones, el tamaño de la etapa decae (debido a que $\beta < 1$), hasta alcanzar Δ_{\min} . Al igual que ADM, Δ_{\min} y Δ_{\max} pueden ser elegidos para proporcionar el rango dinámico deseado y un ruido granular bajo durante la condición de canal desconectado.

El parámetro β controla la velocidad de adaptación. Si β está cercano a la unidad, será la razón de refuerzo y $\Delta(n)$ decae en forma lenta, mientras que si β es mucho menor que la unidad, la adaptación es más rápida. Este esquema de adaptación, puede ser ajustado para cualquiera de las dos formas siguientes:

- ADAPTACIÓN SILÁBICA.
- ADAPTACIÓN INSTANTÁNEA.

Generalmente, este sistema ha sido utilizado en situaciones donde se requiere baja sensibilidad para errores de canal, con necesidades de calidad de voz por debajo de los requisitos para canales de comunicación comercial. En esta situación, el coeficiente del predictor está fijado a un valor considerablemente menor a la unidad, y los parámetros del sistema son ajustados para proporcionar adaptación silábica, debido a que el efecto de los errores del canal desvirtúan la calidad.

El precio que paga un sistema CVSD por su insensibilidad a errores, naturalmente es que la calidad decrece cuando no ocurren errores. Aun sin embargo, su mayor ventaja es que tiene la suficiente flexibilidad para proporcionar un compromiso (trade-off) efectivo entre calidad y la robustez. Todas las técnicas de codificación de forma de onda en el dominio del tiempo, son las más ampliamente utilizadas, están englobadas en lo anteriormente explicado, sin embargo, también hay codificadores de forma de onda en el dominio de la frecuencia.

3.7. CODIFICACIÓN EN SUB-BANDAS (SBC)

Una de las formas más simples de codificación de la forma de onda en el dominio de la frecuencia es la codificación en sub-bandas(SBC), la cual se muestra en el diagrama a bloques en la Fig.3.13. En la aproximación SBC, la señal de voz se subdivide dentro de un conjunto de componentes separados en el dominio de la frecuencia y estos componentes se codifican independientemente. La banda de voz se divide de 4 a 8 sub-bandas, por medio de un banco de filtros, las señales resultantes de estos filtros se codifican usando métodos de cuantización adaptiva (codificadores ADPCM).

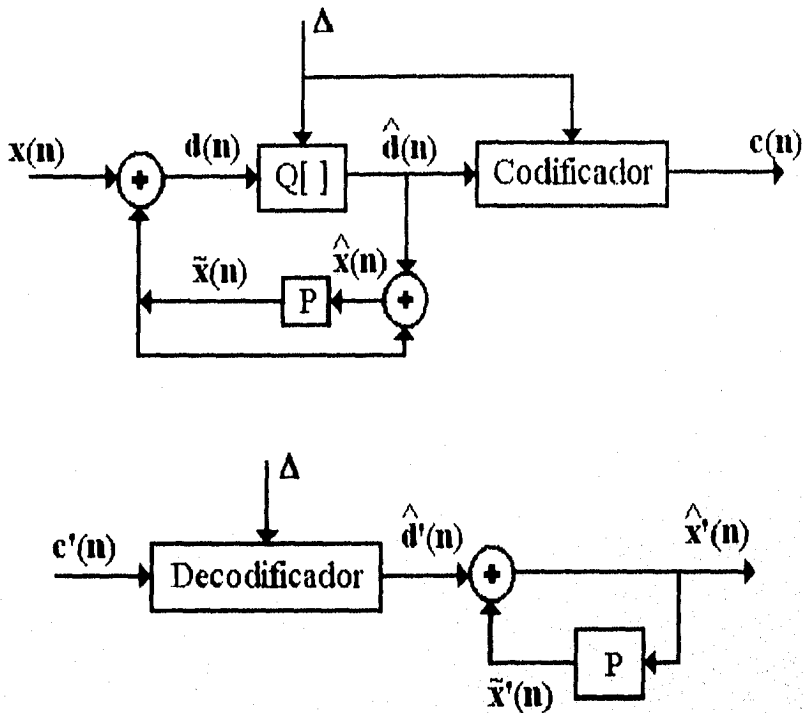


Fig. 3.12.- Diagrama a bloques de esquema de cuantización diferencial: a) Codificador, b) Decodificador.

Debido al tipo de codificador usado, el tamaño de la etapa varía independientemente, las bandas con baja energía usan etapas de tamaño pequeño, produciendo ruido cuantizado menor que con sistemas menos flexibles. Por tanto, el ruido de una banda no afecta otras bandas de frecuencia. Se debe elegir cuidadosamente el número de bits/muestra a utilizarse para codificar cada banda. Después de bajar el rango de muestreo en cada banda, se puede obtener una razón de bit total más pequeña, mientras se mantiene la calidad de la voz, reduciendo el número de bits/muestra en bandas de alta frecuencia que perceptualmente son menos importantes.

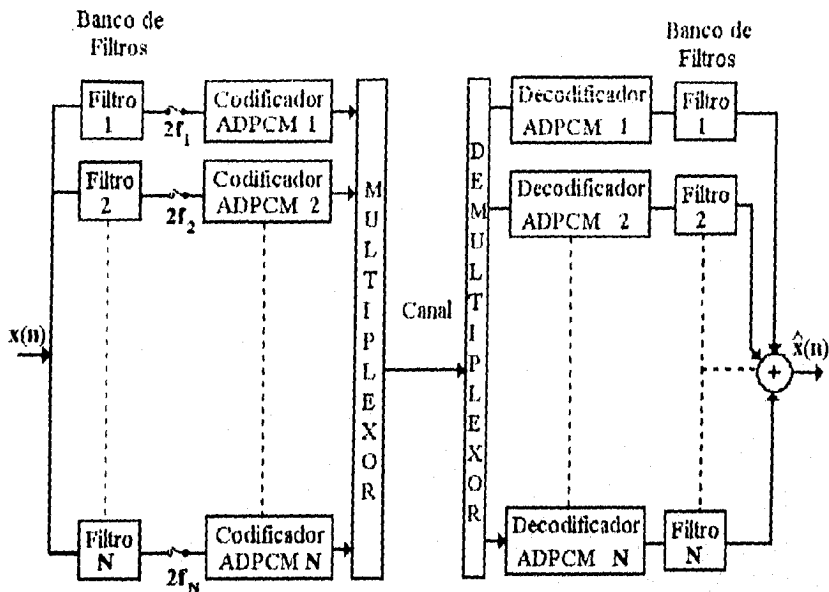


Fig. 3.13. Diagrama a bloques del sistema de Codificación en Sub-Bandas (SBC).

Esta técnica puede adoptar dinámicamente las variaciones espectrales ordinarias (estructura de "formant") en la voz, y proporciona condición de ruido "noise shaping", tomando ventaja de las propiedades de la percepción de la voz (tabla 3.3.1.), y de la insensibilidad de fase del oído. Si se incrementa el número de las bandas de frecuencia y además se permite que el número de bits/muestra varíe dinámicamente de una banda a otra, obtenemos el método de codificación adaptiva por transformada (ATC).

La aproximación ATC, se muestra por medio de un diagrama a bloques en la Fig.3.14. Como se puede apreciar en la figura, a diferencia del sistema SBC, en el sistema ATC las muestras de voz son transformadas a la frecuencia por medio de un banco de filtros y lo realizan efectivamente por medio de un algoritmo de transformada rápida discreta, y entonces los coeficientes de transformación son codificados como un bloque.

En la síntesis se utiliza la información lateral, que afecta las características del cuantizador para transformar los coeficientes retrasados en una señal de voz. Además, esta aproximación es capaz de adoptar dinámicamente y en forma completa las propiedades espectrales de la voz, incluyendo la estructura de "formant", la estructura de tono "pitch", y controla la condición de ruido "noise shaping" por medio de la colocación del bit de asignación o designación en el algoritmo. Tanto la aproximación SBC como la ATC proporcionan una buena calidad de voz a una razón de bit de 16 Kbps.

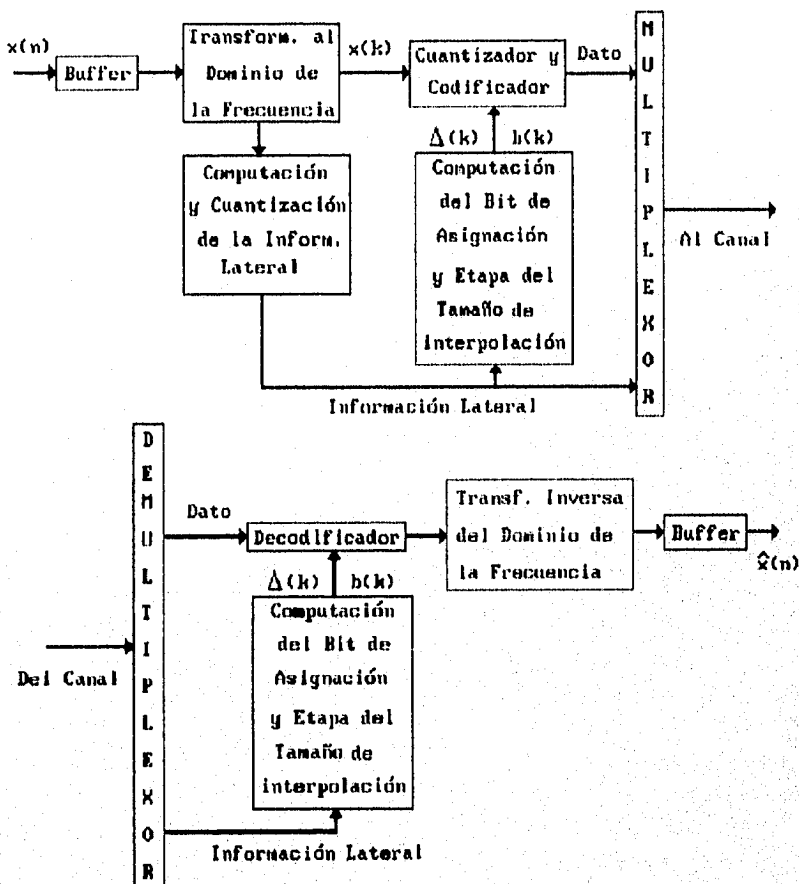


Fig. 3.14.- Diagrama a bloques del sistema de Codificación Adaptiva por Transformada (ATC).

3.8. TÉCNICAS DE MODULACIÓN EN AMPLITUD

Usando esta técnica la señal portadora varía su amplitud de acuerdo a los datos binarios. La técnica digital AM más simple es la DSB (Doble Banda Lateral) o doble banda lateral modulada por una señal binaria. La señal DSB es representada por:

$$f(t) = \frac{A}{2} (1 + m(t)) \cos \omega_c t \quad (3.8.1)$$

AM-DSB

Donde $m(t)$ es la señal modulante y ω_c es la frecuencia en radianes por segundo. Dado que la portadora no contiene información, la eficiencia puede ser mejorada por el uso de una variante con DSB portadora suprimida, así:

$$f(t) = A m(t) \cos \omega_c t \quad (3.8.2)$$

AM-DSB-SC

La situación cuando $m(t) = \pm 1$ en la ec.3.8.1. y $m(t) = 1.0$ en la ec.3.8.2. produce la modulación OOK (on-off-keying) la cual es un tipo de señal que tiene como característica el encendido y apagado en una onda portadora de alta frecuencia. La Fig.3.15. ilustra el procedimiento para obtener este tipo de modulación. Si $m(t) = \pm 1$ en la ec.3.8.2. se obtiene una señal BPSK (binary shift keying), la cual se discutirá al tratar las técnicas de modulación en fase. Para rescatar una señal de este tipo, en el receptor se utiliza un detector no-coherente de los llamados de envolvente, cuyo circuito más conocido es el RC usado en los receptores de radio comercial Fig.3.16.

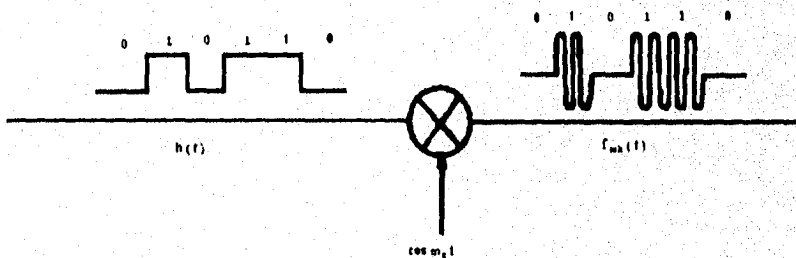


Fig. 3.15. Modulador OOK.

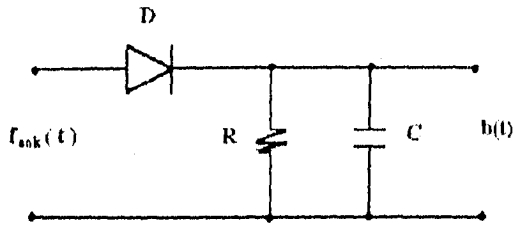


Fig. 3.16. Detector de envolvente.

3.9. TÉCNICAS DE MODULACIÓN EN FRECUENCIA

La técnica básica de modulación en frecuencia es la FSK (frequency shift keying) donde la señal binaria $b(t)$ es usada para generar la forma de onda:

$$f(t) = A \cos(\omega_c \pm \Omega)t \quad (3.9.1.)$$

FSK

En la cual el signo más o menos se aplica dependiendo si el bit es 0 ó 1. La señal transmitida, entonces tiene una frecuencia $\omega_c - \Omega$ o $\omega_c + \Omega$, siendo Ω la desviación de frecuencia angular de la frecuencia portadora ω_c . La Fig.3.17. muestra la forma de onda obtenida con esta técnica.

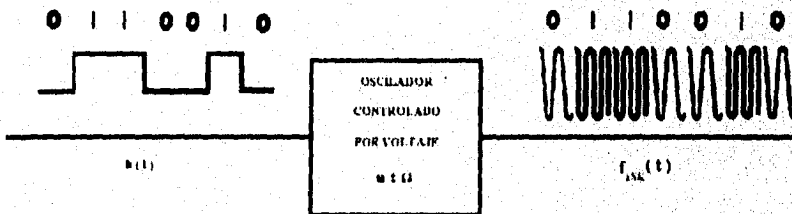


Fig. 3.17. Modulador FSK.

Con esquemas FSK, es práctica común especificar la separación en frecuencias, en términos de un índice de modulación, definido por:

$$\beta = \Delta f T \quad (3.9.2)$$

donde T es la duración del símbolo (igual al inverso de la rapidez de la fuente binaria $b(t)$). Como en otros esquemas de modulación, FSK puede ser detectado de manera coherente con no-coherente. Detección no-coherente puede ser efectuada por dos filtros pasa banda seguidos por detectores envolventes y un dispositivo de decisión. Con esta aproximación la separación de frecuencias debe ser al menos $1/T$ ($\beta \geq 1$) para prevenir significativo traslape en la banda de paso de los dos filtros.

En cuanto a detección coherente la señal podría ser recobrada multiplicando por dos señales, una con frecuencia $\omega_c - \Omega$ y otra con frecuencia $\omega_c + \Omega$, seguidas de filtros pasa bajas, que una vez restadas sus salidas entregan los valores binarios (Fig.3.18).

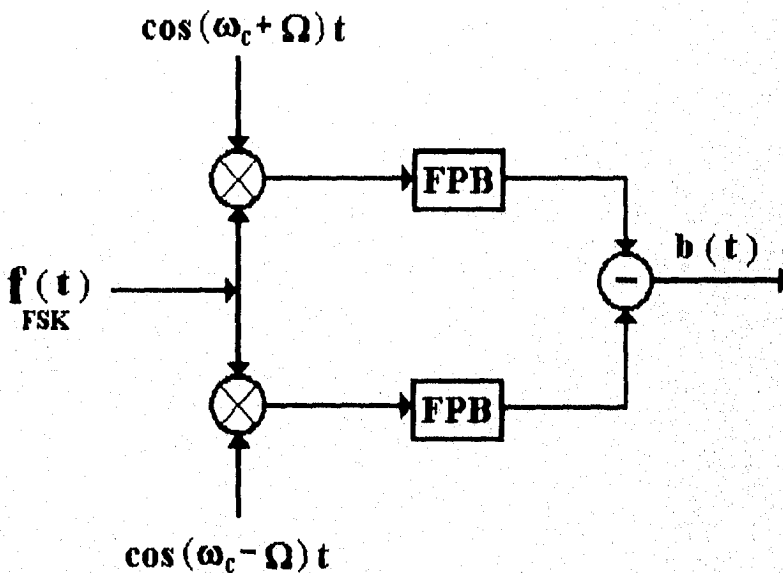


Fig. 3.18. Detección coherente en FSK.

FSK de fase continua (CPFSK) es otro esquema de modulación FSK en el cual la fase es obligada a ser constante durante una transición de símbolo, esto es, los cambios abruptos de fase en los instantes de transición de bits, característicos de otras implementaciones FSK, son evitados. Este detalle de fase resulta en ventajas en la banda espectral de frecuencia así como en mejorada eficiencia por el uso de intervalos de observación mayores de un bit. Con detección coherente, valores de β en las vecindades de 0.7 han mostrado proveer un óptimo funcionamiento para cualquier intervalo de observación. Un caso especial de CPFSK, de gran interés en años recientes, es la técnica MSK (minimum shift keying), a la cual corresponde una β de 0.5, cuya característica principal es la mínima separación en frecuencia de los tonos de señalización, esto se refleja en que hay exactamente medio ciclo de diferencia entre el tono de 1 y el tono de 0, según se observa en la Fig.(3.17.). Es importante hacer notar aquí, que generalmente f_1 y $f_2 > 1/T$. En algunos sistemas, particularmente sobre líneas telefónicas, f_1 y $f_2 = 1/T$, como es el caso de los sistemas CPFSK.

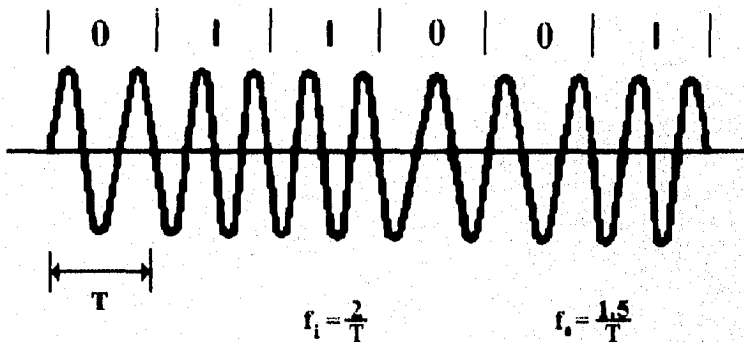


Fig. 3.19. Señal MSK.

3.10. TÉCNICAS DE MODULACIÓN EN FASE

Esta categoría de modulación angular hace uso de cierta separación en fase de una señal portadora con el fin de representar una serie de bits. Nos hemos referido ya a la técnica básica de modulación en fase BPSK en la sección 3.8., lo cual nos da una idea del tipo de señales que obtendremos al estudiar las señales del tipo M-PSK.

PSK (Phase Shift Keying) ha resultado ser la técnica de modulación más popular para densidades de información intermedias y aplicaciones de gran nivel de trabajo. La popularidad es primeramente debido a dos razones, insensibilidad a variaciones de nivel y buena capacidad para evitar transmisión de errores. La expresión de una señal M-PSK está dada por:

$$f_i(t) = A (\cos(\omega_c t + \phi_i)) \quad (3.10.1)$$

PSK

donde: A = amplitud de la señal PSK.

ω_c = frecuencia de la portadora en rad/seg.

ϕ_i = fase variable a datos binarios

Desarrollando la ec. anterior,

$$f_i(t) = A (\cos \phi_i \cos \omega_c t - \sin \phi_i \sin \omega_c t) \quad (3.10.2)$$

PSK

Nos damos cuenta que $\cos \phi_i$ y $\sin \phi_i$ son constantes sobre un intervalo de señalización, y por tanto representan los coeficientes para expresar $\cos(\omega_c t + \phi_i)$ como una combinación lineal de las señales $\cos \omega_c t$ y $\sin \omega_c t$. Es decir:

$$f_i(t) = A (C_i \cos \phi_i \cos \omega_c t - S_i \sin \phi_i \sin \omega_c t) \quad (3.10.3)$$

PSK

donde $C_i = \cos \phi_i$ y $S_i = -\sin \phi_i$

Dado que $\cos \omega_c t$ y $\sin \omega_c t$ están fuera de fase 90° uno respecto a otro, son ortogonales en un diagrama de fase, y por tanto se dice que están en cuadratura. En esencia, $\cos \omega_c t$ y $\sin \omega_c t$ representan vectores base en un diagrama de fasores bidimensional. La señal coseno se dice que está dentro de fase y es conocida como señal I, la señal seno se dice que está fuera de fase y se le denomina señal Q. La tabla 3.10. muestra los coeficientes en cuadratura C_i y S_i necesarios para generar una señal PSK de 2, 4, y 8 fases. En la Fig.(3.20.) se muestra el diagrama de fases de las señales anteriores, donde los puntos que representan a las señales son conocidos como constelación de señales; a un lado aparecen las formas de onda obtenidas para cada intervalo de señalización.

FASE	SEÑAL COMPUESTA	C_i	S_i	VALOR BINARIO
2ϕ	$A \cos (\omega_c t + \Pi / 2)$	0.0	-1.0	0
	$A \cos (\omega_c t - \Pi / 2)$	0.0	1.0	1
4ϕ	$A \cos (\omega_c t + \Pi / 4)$	0.707	-0.707	01
	$A \cos (\omega_c t + 3\Pi / 4)$	-0.707	-0.707	00
	$A \cos (\omega_c t - 3\Pi / 4)$	-0.707	0.707	10
	$A \cos (\omega_c t + \Pi / 4)$	0.707	0.707	11
8ϕ	$A \cos (\omega_c t + \Pi / 8)$	0.924	-0.383	011
	$A \cos (\omega_c t + 3\Pi / 8)$	0.383	-0.924	010
	$A \cos (\omega_c t + 5\Pi / 8)$	-0.383	-0.924	000
	$A \cos (\omega_c t + 7\Pi / 8)$	-0.924	-0.383	001
	$A \cos (\omega_c t - 7\Pi / 8)$	-0.924	0.383	101
	$A \cos (\omega_c t - 5\Pi / 8)$	-0.383	0.924	100
	$A \cos (\omega_c t - 3\Pi / 8)$	0.383	0.924	110
	$A \cos (\omega_c t - \Pi / 8)$	0.924	0.383	111

Tabla 3.10.- Coeficientes en cuadratura para PSK de 2, 4 y 8 fases.

Un método general para producir la modulación en fase y generar señales Mary-PSK basado en los coeficientes en cuadratura, se muestra en la Fig.(3.21.)

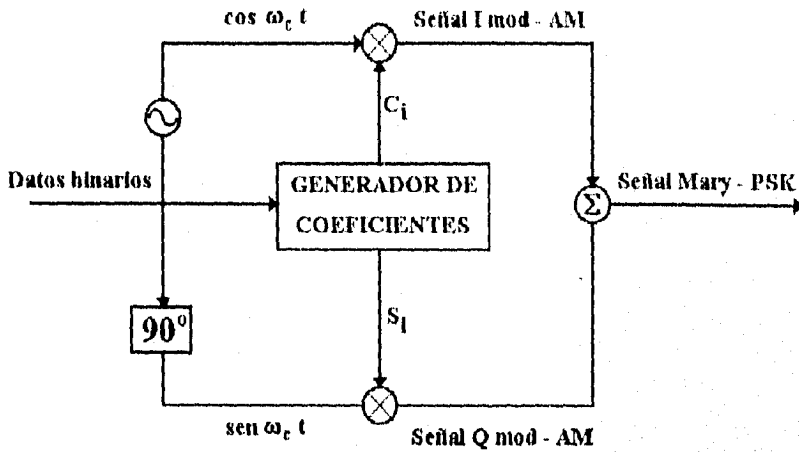


Fig.3.21. Modulador Mary - PSK.

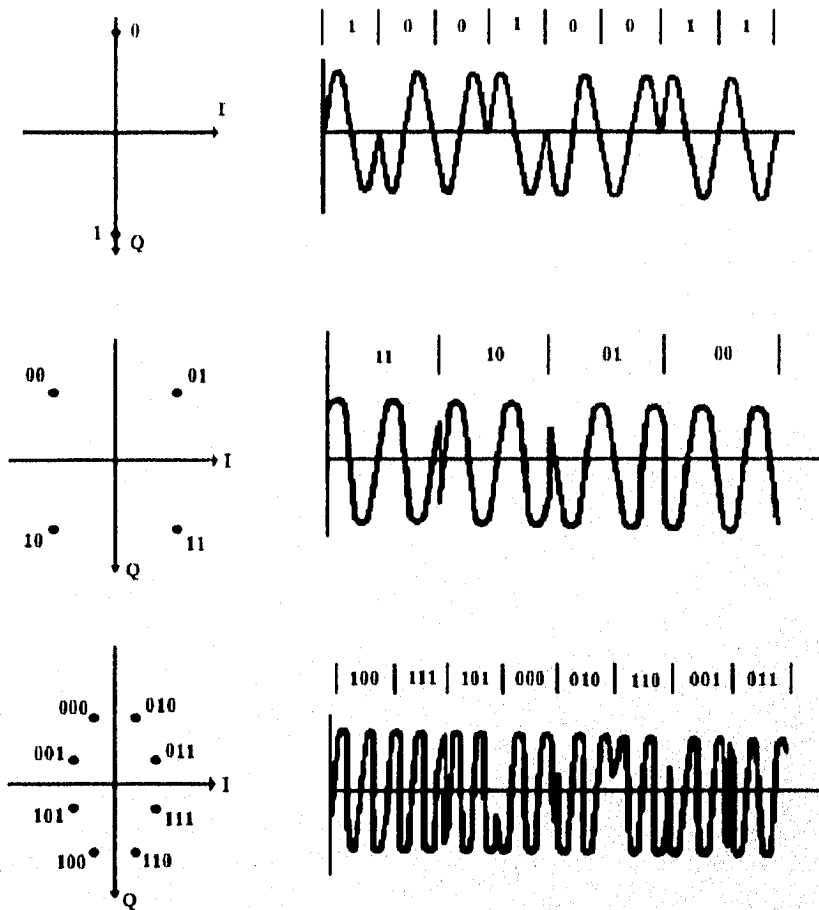


Fig. 3.20. Diagrama fasorial y formas de onda para PSK de 2, 4 y 8 fases.

3.11. TÉCNICAS DE MODULACIÓN HÍBRIDAS

En años recientes la utilización de técnicas híbridas que hacen uso de la combinación de las tres formas de modulación típicas (AM, FM y PM) para representar un tren de datos binarios, ha causado un gran impacto en las redes de telecomunicaciones. La técnica de más utilidad es la conocida como corrimiento de amplitud - fase o APK y cuando es usada en cuadratura se reconoce también con modulación en amplitud de cuadratura QAM. Bajo esta técnica los coeficientes en cuadratura son valuados de modo tal que entregan una fase i y una amplitud multinivel i . De esta forma QAM es bastante semejante a PSK, excepto que se tiene una amplitud o varias para una cierta fase. De acuerdo a las consideraciones anteriores, la señal QAM se puede representar por la siguiente ecuación:

$$f_i(t) = A_i (\cos(\omega_c t + \phi_i)) \quad (3.11.1.)$$

QAM

donde A_i = amplitud multinivel i

ω_c = frecuencia de la portadora en rad / seg.

ϕ_i = fase variable a datos binarios

Siguiendo un desarrollo algebraico, la ec. anterior resulta :

$$f_i(t) = C_i \cos \omega_c t + S_i \sin \omega_c t \quad (3.11.2.)$$

donde $C_i = A_i \cos \phi_i$ y $S_i = -A_i \sin \phi_i$

Aquí, es útil expresar las siguientes igualdades:

$$A_i = \sqrt{C_i^2 + S_i^2} \quad (3.11.3.)$$

$$\phi_i = \arctan \left(\frac{-S_i}{C_i} \right) \quad (3.11.4.)$$

De esta forma podemos conocer la expresión para la señal $f_c(t)$ en base a la constelación de señales indicada en un diagrama de fase dado que son conocidos los coeficientes C_i y S_i . En la Fig.(3.22.) se muestran los diagramas de fase para una señal Mary-APK con $M = 8$ y 16 estados.

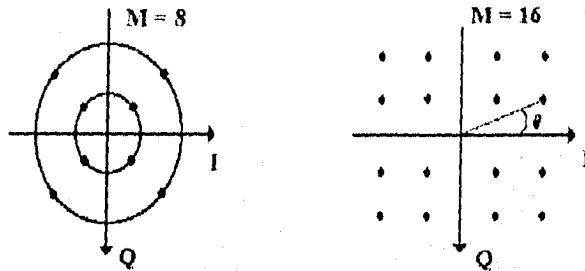


Fig. 3.22. Diagrama fasorial para APK con 8 y 16 estados.

3.12. ELEMENTOS DE JUICIO PARA SELECCIONAR UNA TÉCNICA DE MODULACIÓN

De las anteriores técnicas de modulación, existen factores que favorecen unas sobre otras, la selección se realiza de acuerdo a la aplicación específica. Existen las características espectrales tales como el ancho de banda requerido para transmitir a una específica rapidez de transmisión; tal como la extensión a la cual la señal puede interferir con otros canales adyacentes, medida por la atenuación del espectro en potencia de la señal a una distancia especificada de la frecuencia central. Otro factor son los efectos de interferencia causados por canales adyacentes que acarean con ello cierta cantidad de degradación en la señal modulada. El efecto del desvanecimiento (fading) causado por componentes multiruta (la interferencia CW puede representar la señal de la ruta secundaria). El efecto de distorsión por retraso, el cual es causado por el sistema y no por el canal. Finalmente, el costo y complejidad del esquema de modulación debe ser tomado en cuenta.

3.13. EL SISTEMA TRANSMISOR - RECEPTOR

La Fig.(3.23.) muestra un diagrama a bloques del sistema digital de transmisión. La función del transmisor consiste en codificar los datos entrantes, modulación, modelado del pulso y amplificación de potencia. El codificador convierte los datos de entrada al receptor (polar, bipolar, etc.) a señales entendibles por el modulador, el cual desarrolla ondas portadoras que son modeladas por la subsecuente red de modelado. Este modelado tiene por objeto suprimir la energía de la señal fuera del ancho de banda del canal asignado. En el receptor otra red de modelado es diseñada para rehacer la forma al pulso transmitido, es decir, para rechazar el ruido y suprimir la interferencia de canales adyacentes.

La señal resultante de esta etapa es muestreada a instantes de tiempo apropiados y decisiones en la información son hechas por el demodulador y decodificador. En el receptor referencias de tiempo y frecuencia deben ser tomadas de la señal transmitida. Errores en las referencias recobradas resulta en un funcionamiento no sincronizado en el sistema. La sensibilidad de varias formas de modulación y modelado de pulsos a errores en las referencias locales es un factor importante en la selección de un sistemas de transmisión.

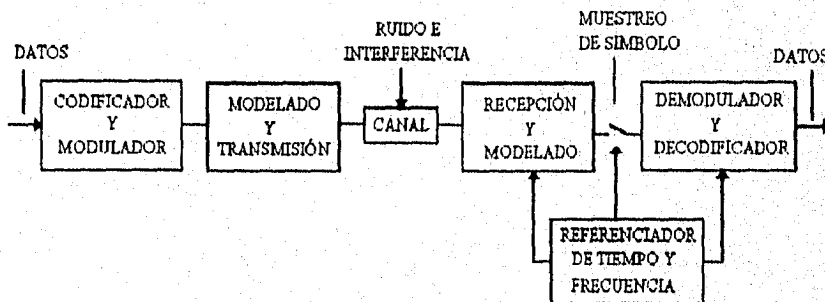


Fig. 3.23. Sistema transmisor - receptor por onda portadora.

CAPITULO 4

REPORTE DE APLICACIÓN DEL TMS3477

CONSIDERACIONES

El TMS3477 CVSD (DeltaModulación con Pendiente que Varía Continuamente) es un digitalizador de señales analógicas que provee a bajo costo una solución para muchas aplicaciones que requieren precisión, exactitud y reproducción en la grabación de señales. El TMS3477 fué diseñado con interface directa a DRAM'S (Memorias de Acceso Aleatorio Dinámico) e incluye todo lo necesario para operar como todo un circuito en un solo chip. Sin embargo, muchas aplicaciones requieren almacenar datos en otros dispositivos para distribuirlos en estaciones de reproducción. Esto es, se requiere una interface a memorias no volátiles o dispositivos que almacenen dicha información ya que los datos en una DRAM no son fáciles de transferir. En este capítulo aplicaremos el TMS3477 para estas situaciones. Para lo antes mencionado es necesario crear dos circuitos que mantienen una relación muy estrecha con el digitalizador y son :

1) La Estación de Grabación.- Ésta captura y reproduce el flujo de datos usando una memoria no volátil SRAM (Memoria de Acceso Aleatorio Estática) y a travez de una carga directa de estos datos a una PROM (Memoria de Solo Lectura Programable).

2) La Estación de Reproducción.- Reproduce los datos programados en una PROM.

INTRODUCCIÓN

El TMS3477 es un simple chip que controla la grabación / reproducción de señales análogicas y la realiza a traves de Modulación Delta con Pendiente que Varía Continuamente (CVSD) Fig.4.1. La idea básica del circuito es la transformación por muestreo de una señal analógica en digital para su grabación en un dispositivo y despues la transformación inversa de la señal digital en analógica con la lectura de los datos almacenados. Esta señal es entonces convertida en un valor digital que puede ser grabado. Para que la señal tenga una cierta fidelidad es preciso que cada ciclo tenga una cantidad mínima de muestras. De esta forma, cuanto mayor sea la frecuencia con que se toman las muestras, mayor será la fidelidad en la grabación y en la reproducción. Sin embargo aumentado la frecuencia ganamos fidelidad, pero disminuimos el tiempo de grabación. La frecuencia de muestreo es seleccionada por el TMS3477 entre 16, 32 y 64 kHz Fig 4.1.1.

Este dispositivo convierte datos análogos en una representación de datos digitales comprimidos y supervisa la administración de memoria de estos datos dentro de una DRAM. Los comandos de GRABACIÓN, REPRODUCCIÓN, PAUSA y ALTO directamente se pueden interpretar como los usados en un grabador de cinta convencional y pueden se accesados a travez de una tecla o modo de un microcontrolador. El hardware del TMS3477 puede programarse en diferentes configuraciones, ademas opera con tres diferentes tamaños de memorias. Puede grabar y reproducir una o dos frases, operar hasta con más de cuatro chips de memoria (no importa el tamaño) y grabar un tamaño de frase fija o variable (para aplicaciones de voz).

Graba la longitud de un señal fija o reproduce continuamente la información importante anteriormente almacenada; puede reproducir al mismo tiempo la señal que entra a la grabación. Como el TMS3477 ofrece un programador de señales con un sistema de alta fidelidad; lo aplicaremos a circuitos de bajo costo que son sencillos de operar y programar. Esto extiende la aplicación del TMS3477 en áreas que requieren de ROM (Memorias de Sólo Lectura) con una rápida y simple reproducción de las señales. Para la reproducción, el circuito lee en secuencia los valores digitales almacenados y los convierte nuevamente en tensión (esta tensión de salida se realiza por medio de Modulación por Amplitud de Pulso PAM) y fluye por un integrador de modo que haya un "filtrado" que permita recomponer la señal original Fig.4.1.2. En este capítulo se diseño el circuito, y se explica la circuiteria y la metodología necesaria en la captura del código digital de datos CVSD, donde el TMS3477 será la estación base para la distribución y reproducción de las señales almacenadas. Particularmente se describe la circuiteria con la manipulación de la fuente de datos de entrada y salida del TMS3477 para utilizarse con un direccionamiento de byte o palabra de 8 bit en SRAM'S (Memorias de Acceso Aleatorio Estatica) y EPROM'S (Memorias de Solo Lectura Programable y Borrable) o otro dispositivo magnético (cinta, disco, etc.).

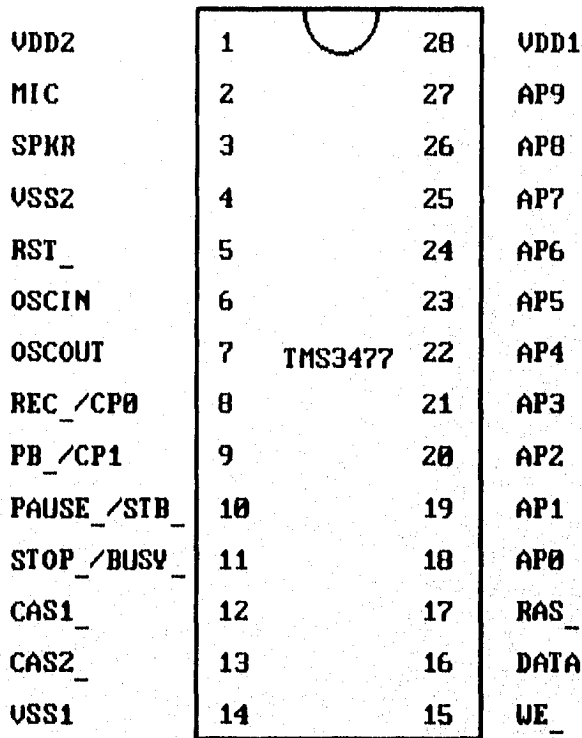


Fig.4.1. Diagrama de configuración del TMS3477.

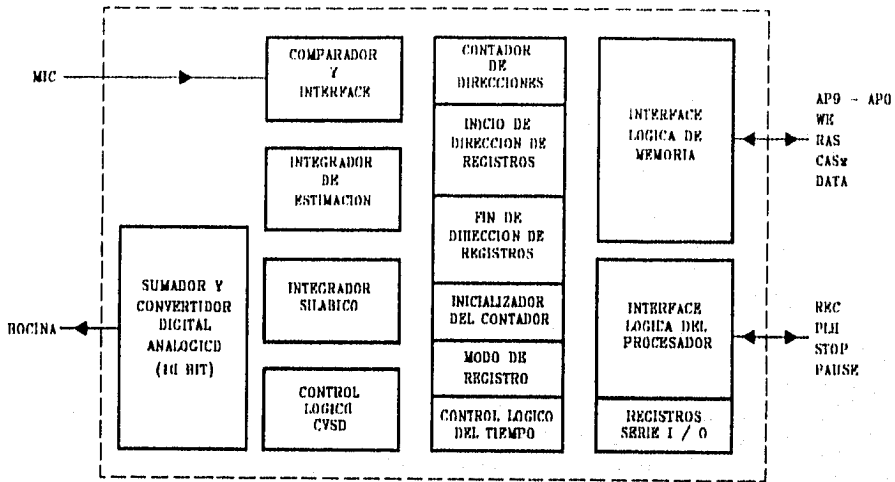


Fig.4.1.1. Diagrama de Bloques del TMS3477.

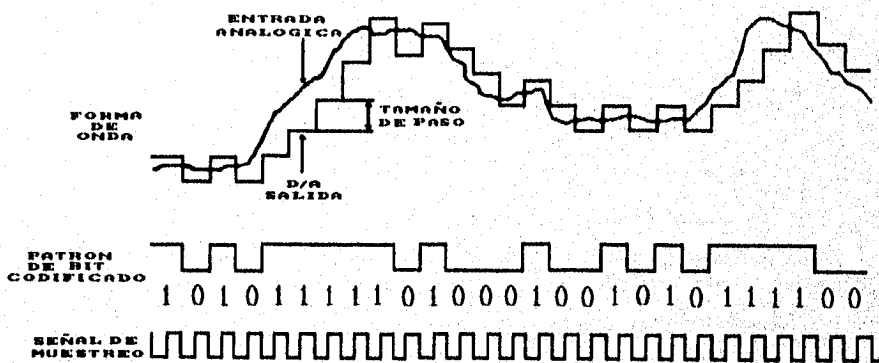


Fig.4.1.2. Modulación CVSD.

4.1. FUNCIONAMIENTO DEL CIRCUITO

La aplicación del diseño consiste en dos circuitos : estación de Grabación/ Reproducción y estación de Reproducción. La estación de Grabación / Reproducción desarrolla la señal digitalizada. Los datos de la SRAM son transferidos a la estación de memoria PROM conectada vía bus de datos en el circuito de grabación/reproducción.

En la estación de Reproducción los datos son usados para programar ROM'S y poder realizar la reproducción de la señal en varios circuitos.

4.1.1. Estación de Grabación / Reproducción

Este circuito funciona como una estación que reproduce señales. Acepta entradas analógicas (como un micrófono) y genera señales que pueden excitar una bocina y reproducir la señal almacenada con alta fidelidad. Tiene cinco funciones :

- GRABACIÓN
- REPRODUCCIÓN
- PAUSA
- ALTO
- INICIALIZACIÓN

Generalmente estas funciones trabajan cuando son solicitadas por el operador, sin embargo algunas veces existen diferencias cuando se esta grabando una señal. El TMS3477 graba el comienzo y el final de una(s) señal(es) a destinar en una registradora que no es accesible directamente de la circuitería exterior. El TMS3477 tiene un sistema de configuración con opción para fijar o variar el tiempo de muestreo y grabación. Cuando no se fija parámetro alguno al operar el TMS3477 inmediatamente realiza sus funciones de forma normal esto es : da comienzo a la grabación de la señal de entrada, envía datos a almacenar y da término de la señal a grabar. Sin embargo si se inicializa o se interrumpe la energía estos registros internos son puestos a cero (desde que el TMS3477 es energizado manda directamente cada evento y este es borrado por la memoria).

Aún así si una grabación fija reside en memoria no volátil, el TMS3477 no puede reproducirla en el modo de longitud variable, porque este sólo puede intentar reproducirlo entre las direcciones internas almacenadas. Por esta razón, el modo de longitud fija de grabación es preferible para este circuito. En este modo la función de ALTO limpia el contador de direcciones a cero y la función de REPRODUCCIÓN reproducirá totalmente lo almacenado en memoria. De esta manera la señal final grabada puede consumir totalmente la memoria o le incluye silencio desde el final de la señal hasta que se agoten las direcciones de almacenamiento de la memoria. Por otra parte, en REPRODUCCIÓN, es usualmente desagradable e impredecible tener ruidos (Térmico, Ambiental, Sobremuestreo , etc.) que pueden seguir a la señal grabada. O, si una señal pequeña es grabada

sobre una señal previa más grande, la reproducción incluirá la parte final de la primera señal que no fué sobrescrita. Siempre que la función de ALTO se selecciona, inicializa el contador de direcciones a cero y debe reproducirse el mensaje nuevamente. Si se selecciona la función de PAUSA simula una suspensión, y puede continuar la reproducción a partir de la dirección donde se interrumpió la señal. Esta estación también tiene la capacidad de transferir los datos almacenados en memoria a una unidad PROM via bus de datos y direcciones.

4.1.2. Estación de Reproducción

Este circuito tiene la misma funcionalidad que en la estación de Grabación/Reproducción de la sección anterior. Pero no tiene la circuitería de entrada de la señal analógica ni tampoco la circuitería de multiplexaje de escritura en memoria. Sólo tiene las funciones de:

- REPRODUCCIÓN
- ALTO
- PAUSA
- INICIALIZACIÓN

y la salida de la señal hacia un amplificador. Este circuito está proyectado para ser una estación de Reproducción de señales analógicas localizada remotamente lo más barata posible. Utiliza dos circuitos integrados más y un transistor que realiza las funciones de RAS (Selector de Dirección de Renglón) en el circuito de la sección anterior.

4.2. Operación del circuito.

DIGITAL

Los puntos de primordial importancia en el manejo de datos digitales de este sistema son : el esquema de conversión de direcciones, el método de multiplexaje de datos, los detalles de las memorias de lectura/escritura, la sección de configuración de resistores y la operación de la interface de carga de datos.

4.2.1. Conversión de Direcciones

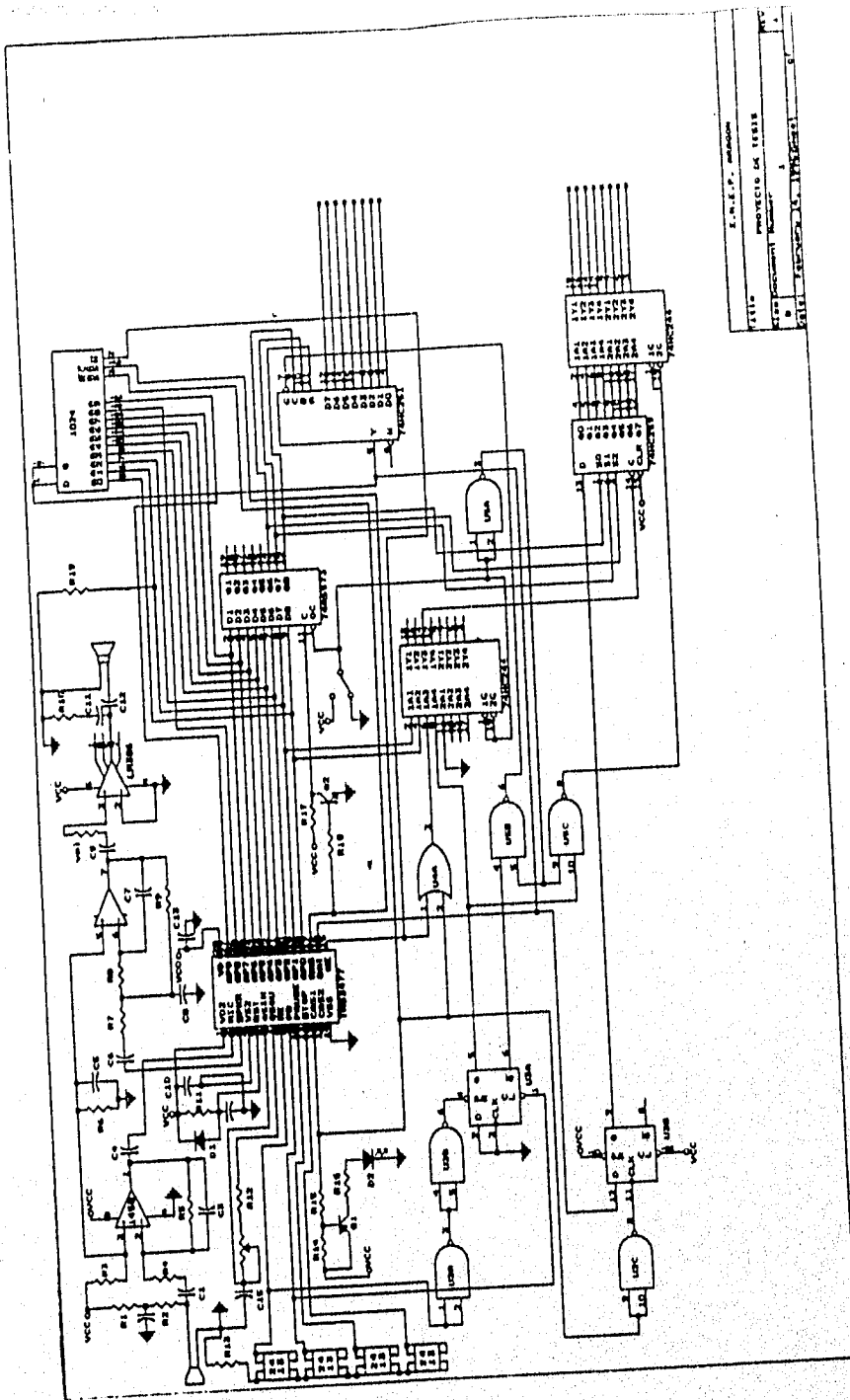
El diseño del TMS3477 se realizó pensando en el uso de DRAM'S, y por tanto el modo de direccionamiento de datos es de acuerdo al protocolo de renglón y columna de las memorias dinámicas. Para realizar la grabación o la transferencia eventual de datos a memoria en una EPROM, el almacenaje debe ser hecho de una memoria con un esquema similar de dirección de byte. Como esto no es posible, una solución sería una base de memoria SRAM de 128K x 8. Sin embargo, las SRAM'S y DRAM'S usan diferentes métodos de direccionamiento. Las SRAM requieren todas las líneas de direcciones al mismo tiempo, mientras que en las DRAM sus segmentos de dirección deben ser separados por un selector de renglón y columna. Por tal motivo al diseñarse el circuito se pensó en realizar una interface para esta manipulación de direcciones de la DRAM con el TMS3477 y la SRAM. Esto se puede simplificar con el circuito integrado 74HC573 que es un latch de bits Fig. 4.2.1.

Este dispositivo recoge los 8 bits más bajos de las direcciones actuales en el borde activo de la señal del RAS (Selector de Dirección de Renglón). Entonces las salidas de las direcciones de columnas del TMS3477 se dedican a llenar las direcciones de memoria en sus terminales. La presencia de todas las direcciones activas se detecta cuando el TMS3477 activa el CAS (Selector de Dirección de Columna). Los tres bits más bajos tomados por las direcciones de renglón, (en este caso los bit Q8, Q7 y Q6 del U6-74HC573) son utilizados para el control del multiplexaje de datos de 1 bit del flujo de datos del TMS3477 en la organización del tamaño del byte de la SRAM.

4.2.2. Multiplexaje de Datos

El tamaño de bit del flujo de datos generado y recibido por el TMS3477 es convertido a un formato de tamaño de byte por el U9-74HC251 (selector de datos y multiplexor) y el U8-74HC259 (direccionador de 8 bit). Las operaciones de ambos multiplexores son controladas por los tres más bajos bits de direcciones del TMS3477 recogidos por el U6-74HC573.

La escritura de la SRAM es realizada usando un Flip-Flop tipo D (U3B-74HC74) como un asignador direccionable, y un manejador de bus de tres estados. Durante el ciclo de escritura el actual bit del (U12-TMS3477), el cual fué recogido por un Flip-Flop tipo D (U3B-74HC74), es asignado dentro de 1 bit en una salida del tamaño de un byte por el (U8-74HC259). La posición particular del bit es controlada, como si fuera un ciclo de escritura, por los tres más bajos bits de direcciones asignados por el (U6-74HC573).



E. R. L. P. M. M. M. M.
 PROYECTO DE TESIS
 N.º de Proyecto: _____
 N.º de Tesis: _____

4.2.3. Memoria Lectura/Escritura

La estación de sólo reproducción no requiere de circuitos para supervisar la memoria de lectura/escritura. Esto lo realiza a través de un bus de datos permanente a la memoria. Por consiguiente no requiere la supervisión del bus de transferencia de datos de la memoria como ya se menciona en la sección anterior. La realización de la estación grabación/reproducción, sin embargo, requiere una circuitería similar. El control de la SRAM es realizado por un flip-flop tipo D (U3A-74HC74), una compuerta OR (U4A-74HC32) y dos compuertas AND (U2A,B-74HC00). Estas partes actúan en dos funciones primarias: bus de control y escritura de SRAM. Esta operación descrita es aproximadamente tan grande como tenga la capacidad de almacenamiento la memoria seleccionada en la estación de grabación/reproducción. La asignación de carga tri-estado (U7-74HC244) asegura que no se perderan datos en el transcurso del almacenamiento a memoria.

Las salidas del flip-flop D (U3A-74HC74) tienen el control maestro de las líneas de datos en la SRAM y la línea de flujo de datos del TMS3477. Las funciones de INICIALIZAR y GRABACIÓN son sumadas a través de (U2B-74HC00), entonces invierte la entrada de inicio que tiene el flip-flop D (U3A-74HC74). Así de esta manera al seleccionar la función grabación, el flip-flop D será inicializado a una salida en alto de manera que establezca una condición de escritura para lo cual el (U12-TMS3477) conduce la línea de flujo de datos y el tri-estado (U10-74HC244) es activado tomando el bus de control de la SRAM al (U8-74HC259). Inverso a esto, al activar la función de reproducción limpia la entrada del flip-flop D (U3A-74HC74) y lo envía a un tri-estado (U10-74HC244) y le permite a la SRAM manejar este bus de datos y (U9-74HC251) selecciona el manejo de la línea de flujo de datos hacia el (U12-TMS3477).

Leer la SRAM es tan fácil como escribirla. No requiere de un circuito de control adicional; líneas superiores que seleccionan el byte y inferiores que seleccionan el bit a través del (U9-74HC251), cuya salida tri-estado se activa para manejar la línea de flujo de datos descrito anteriormente. Esta condición de reproducción se da el lujo de ignorar el intermitente acceso de renovación de la DRAM generado por el (U12-TMS3477) ya que este tiene la inteligencia de distinguir entre una renovación y una dirección real de lectura. Este privilegio del chip por simplicidad permite que ambos sean procesados similarmente. Es decir cualquier acontecimiento tiene que estar fijo para ser tomado por el (U6-74HC573), el cual puede por consiguiente redireccionar la SRAM y el multiplexor. Pero siempre que el (U12-TMS3477) no lee efectivamente una línea de datos durante una actualización, el acceso de direcciones, no requiere un hardware discriminador para eliminar el dato engañoso de la línea de datos del (U12-TMS3477).

El ciclo de escritura requiere de más hardware a causa del tiempo de unión entre el (U12-TMS3477) y la SRAM. Generalmente el trabajo de escribir como tal, es : cuando un bit real está presente en la línea de datos, este es tomado por el propio bit del (U8-74HC259) seleccionado de acuerdo, por los 3 más bajos bits de direcciones tomados del (U6-74HC573) como byte de salida y lo escribe dentro de la memoria. Debe tomarse en cuenta que esta señal escrita es generada por cada bit real, con lo que se pretende que cada byte de memoria sea escrito en ocho tiempos. Cada vez que ocurre una

escritura un bit mas ha sido actualizado, hasta que esten completos los 8. Esto es bastante repetitivo, pero simplifica los requerimientos de hardware. El CAS (Selector de Dirección de Columna) controla el bit de datos tomados y la compuerta OR (U4A-74HC32) maneja la señal de escritura de la SRAM y sirve para fucionar las características de tiempo entre la SRAM y el (U12-TMS3477).

El flip-flop D (U3B-74HC74) mantiene el bit actual disponible para el (U8-74HC259) en el ciclo entero antes del próximo bit real que es señalizado por CAS. La compuerta OR (U4A-74HC32) maneja las señales incluídas; la señal WE del (U12 TMS3477) la cual es retrasada por la señal CAS (Selector de Dirección de Columna). Este retraso proporciona una señal a la compuerta OR (U4A-74HC32) la cual es cortada por el bit que mantiene el flip-flop D (U3B-74HC74). Este arreglo lo garantiza el (U8-74HC259), el cual activa el nivel antes del borde del disparo, siempre que los datos capturados sean correctos. Este corte de señal a la compuerta OR (U4A-74HC32) ademas cumple con el ciclo de requerimiento de escritura para mantener validas las direcciones durante la señal activa de WE en la SRAM. La señal (esta señal se activa primeramente antes que la SRAM y WE) de CAS permanece activa despues de validar la dirección enviada. La compuerta OR (U4A-74HC32) proporciona la interface entre estos dos.

4.2.4. Organización de Resistencias

El TMS3477 en el modo de selección de las líneas de dirección, describe los cambios que tendrán que hacerse para cada diseño en particular. Toda la selección de resistencias son reconocidas por el TMS3477 solo depues de una INICIALIZACIÓN o despues de energizar el circuito. En ambas estaciones la de Reproducción/Grabación y solo Reproducción, puede haber resistencias. El resistor en A3 selecciona la inhibición del fin de direcciones de la opción de longitud fija de frase. El resistor en A1 cambia el tamaño de la memoria a 1Megabit. Los resitores A7 y A6 organizan el modo de control de la frecuencia interna de muestreo y pueden ser seleccionados en ambas estaciones. Si no existe ningún resistor, la velocidad es de 32 Kbit/seg, si existe A7, la velocidad es de 64 Kbit/seg. Si existe A6, la velocidad es de 16 Kbit/seg. Ambos resistores no pueden estar presentes y su comportamiento no esta definido. A9 también puede estar presente en la estación de Reproducción/Grabación, si se selecciona, se puede monitorear por la salida del TMS3477 la entrada que se esta grabando. Esto puede usarse para verificar la propia operación del circuito y comparar las diferentes frecuencias de muestreo. Ver Tabla 4.2.4.

PROGRAMACION DEL MODO DE EJECUCION

RESISTORES EXTERNOS AP9-AP0

9	8	7	6	5	4	3	2	1	0	MODO DE EJECUCION DEL TMS3477
X	X	X	X	X	X	X	X	-	-	DRAM : 256K
X	X	X	X	X	X	X	X	-	R	DRAM : 164K
X	X	X	X	X	X	X	X	R	-	DRAM : 1M
X	X	X	X	X	X	X	X	R	R	RESERVADA
X	X	X	X	X	X	X	-	X	X	CONFIG. DE FRASE : 1 FRASE/ 1 DRAM
X	X	X	X	X	X	X	R	X	X	CONFIG. DE FRASE : 1 FRASE/ 2 DRAMs
X	X	X	X	X	X	-	X	X	X	FIN DE DIRECC. : ACTIVO (FRASE VARIABLE)
X	X	X	X	X	X	R	X	X	X	FIN DE DIRECC. : INACTIVO (FRASE FIJA)
X	X	X	X	X	-	X	X	X	X	CICLO DE GRABACION : INACTIVO
X	X	X	X	X	R	X	X	X	X	CICLO DE GRABACION : ACTIVO
X	X	X	X	-	X	X	X	X	X	MODO MANUAL
X	X	X	X	R	X	X	X	X	X	MODO CPU
X	X	-	-	X	X	X	X	X	X	RELOJ DE MUESTREO(Fbds) = 32 KHz
X	X	-	R	X	X	X	X	X	X	RELOJ DE MUESTREO(Fbds) = 16 KHz
X	X	R	-	X	X	X	X	X	X	RELOJ DE MUESTREO(Fbds) = 64 KHz
X	X	R	R	X	X	X	X	X	X	RESERVADA
X	-	X	X	X	X	X	X	X	X	COMPRESION DE DATOS : INACTIVO
X	R	X	X	X	X	X	X	X	X	COMPRESION DE DATOS : ACTIVO
-	X	X	X	X	X	X	X	X	X	MONITOR DE GRABACION : INACTIVO
R	X	X	X	X	X	X	X	X	X	MONITOR DE GRABACION : ACTIVO

R : RESISTOR EXTERNO EN APn
 - : SIN RESISTOR EXTERNO EN APn
 X : NO IMPORTA

Tabla 4.2.4. Organización de Resistencias.

ANALÓGICA

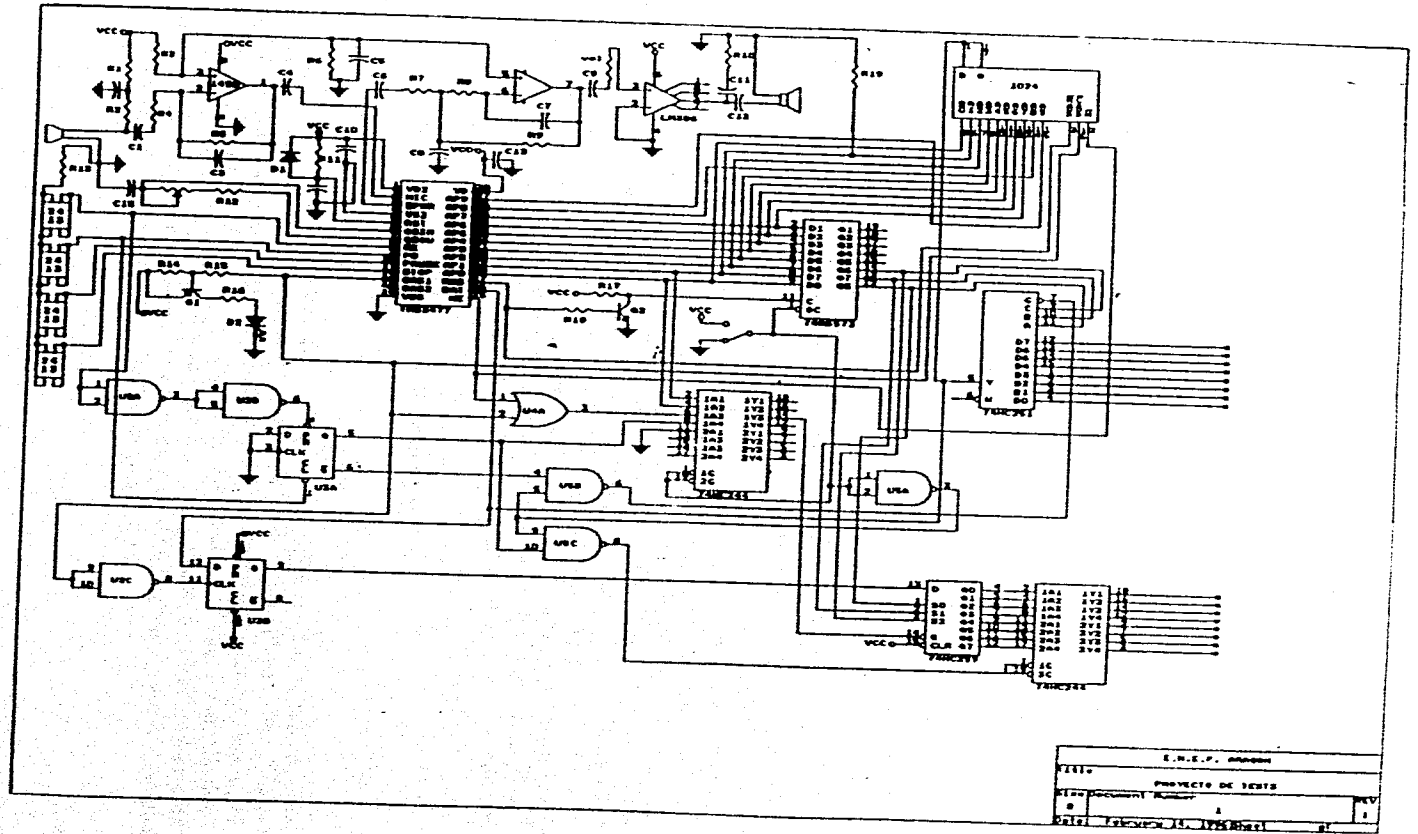
Los principales elementos de operación analógica de este sistema son: el amplificador de la línea de entrada y filtro de audio (U1 LM358), y el amplificador de audio (U LM386) en la línea de salida, la bocina, el micrófono y el oscilador.

4.3.1. La línea de Entrada

Un diagrama esquemático del circuito completo es mostrado en la Fig. 4.3.1. Un micrófono electret que se describirá más adelante es el elemento piezoeléctrico o transductor (el cual tiene una respuesta horizontal de frecuencia menor de 3 dB abajo de 20 Hz) usado para la señal de entrada analógica en el TMS3477 que en este caso es la voz humana. El micrófono, conectado, convierte las ondas sonoras generadas por la voz en impulsos eléctricos, la que es enviada a una entrada inversora del U1 para su amplificación (medio OPAM dual LM358). Obviamente este circuito no funciona en la estación de reproducción. La entrada no inversora es alimentada con $[R6 / (R3 + R6)] * V_{cc}$, o $(47/39) * 5V$ que aproximadamente es +2.5 Volt para cumplir con los requerimientos de la línea de entrada del TMS3477. El resto de la configuración es un amplificador configurado como inversor con R4 y R5 que determinan la ganancia del sistema que es alrededor de 40 dB. La función del capacitor C3 realiza la función de filtro y deja fuera las frecuencias superiores a la frecuencia de corte de la voz humana que es aproximadamente de 4.4 KHz además de ayudar a prevenir oscilaciones indeseadas en el amplificador, particularmente con ganancias muy altas. Los capacitores C2 y C4 proporcionan el desacoplo de DC del amplificador. Todo esto en su conjunto forma un filtro/amplificador de audio.

4.3.2. La línea de Salida

La señal a reproducir es obtenida a partir del pin 3 del TMS3477 y aplicada al (U LM358) que opera como un preamplificador Fig. 4.3.1. Esta sección filtra la salida por Modulación de Amplitud de Pulsos (PAM) y aplicada a la entrada no inversora del (U LM358) tenemos una red integradora que consta de un filtro de segundo orden pasabajos con los componente C6, C7, C8, R7, R8, R9. Para este filtro la frecuencia de corte es alrededor de 13.4 KHz por cuestiones de diseño se reparte el circuito utilizado en el amplificador de la línea de entrada. De este integrado, la señal es aplicada al amplificador de corriente (U5 LM386 amplificador de audio) suministra la potencia necesaria para manejar la bocina. La salida es internamente manejada con 2.5 Volt que es la única manera para su operación y entonces requiere un capacitor desacoplador antes de la bocina. Grandes valores de capacitancias permiten el libre paso de bajas frecuencias en el rango de audio. Los componentes RC



E.N.E.P. - 1980/81	
TITULO: PROYECTO DE SISTEMAS	
NUMERO DE DOCUMENTO: 1	PLANO: 1
FECHA: FEBRUARY 15, 1980/81	

a la salida del LM386 a tierra previenen una conducción (de RUIDO) descrita por National Semiconductor como "La parte Baja Borrosa". Estas son altas frecuencias que oscilan de (5 - 10 MHz) que pueden ocurrir durante el swing negativo dentro de una carga que conduce alta corriente. Para mas detalle consultar literatura del dispositivo de National Semiconductor.

4.3.3. La Bocina

Una bocina común, es usualmente un disco cónico poco profundo, o diáfragma, este vibra hacia atrás y hacia adelante en su lugar en sintonía con la señal de audio cuando esta pasa (Una señal de audio es básicamente una señal compleja de AC). Estas vibraciones reproducen sonido al enviar ondas de presión-sonido a través del aire y estas varían como la señal de audio original varía. Las ondas de sonido alcanzan el tímpano (el cual es conectado al cerebro) este traduce las ondas en sonido que para el hombre son comprensibles. Ahora es tiempo de describir como es hecho el diagrama de la bocina para moverse hacia atrás y hacia adelante. Todo mundo conoce como trabaja un electroimán, y como estos magnetos pueden ser hechos para atraer o repeler a otro cuerpo. Se puede hacer un electroimán usando cable blindado y un clavo, y saber cuando pasa la corriente a través del cable en una dirección si se enrolla el cable al clavo y un extremo del clavo es atraído por un polo del magneto mientras que el otro extremo es repelido; si se invierte el flujo de la corriente se produce lo opuesto al clavo. La Fig.4.3.3. muestra un corte vertical de una bocina. Se puede ver lo que se llama cable móvil; un cilindro hueco donde se enrolla el cable, pero no está adherido, como un magneto fijo. Cuando se mueve es llamado diafragma, cuando una señal de audio o AC pasa a través del cable móvil, entonces se simula la atracción y repulsión de un magneto fijo (igual que el electroimán hecho con el cable y el clavo). El magneto fijo es estacionario pero, el diagrama es forzado a moverse hacia atrás y adelante igual que las oscilaciones de la señal de audio. Como el diagrama oscila, este crea ondas de presión o sonido que son detectadas por el oído humano.

4.3.4. El Micrófono

Los micrófonos de electret son transductores que convierten los sonidos en señales eléctricas, funcionando según el principio de variación de la capacitancia entre dos placas. Una placa es fija, y la otra, hecha de un material muy fino; y puede vibrar con el sonido que recibe. Al vibrar, varía su distancia en relación con la placa fija y por consiguiente, la capacitancia existente entre las dos, como se muestra en la Fig.4.3.4. Las variaciones de la capacitancia, aplicadas a un transistor de efecto de campo FET hace modificar la corriente en este componente, obteniéndose con eso en su salida, una señal eléctrica que es amplificada como se indica en el punto 4.3.1.

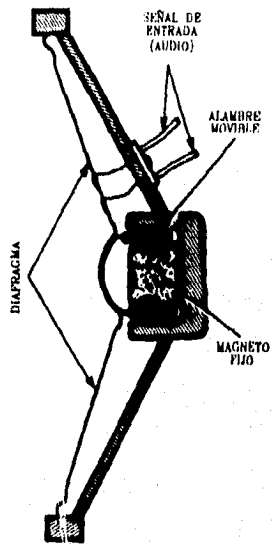


Fig.4.3.3. Corte transversal de una bocina.

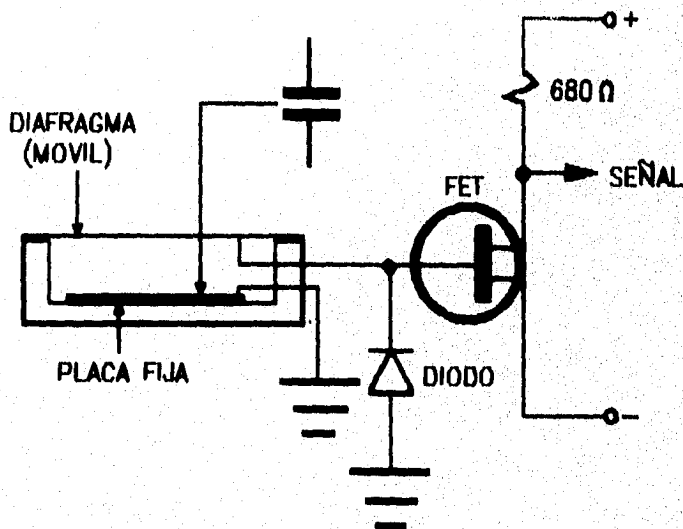


Fig.4.3.4. Micrófono de electret.

4.3.5. El Oscilador

Los sonidos subjetivos de la voz son altamente dependientes en el tono, o frecuencia, de quien produce esos sonidos. Para mantener la integridad de los mensajes almacenados en la SRAM en la estación de Reproducción/Grabación, el oscilador y ambas estaciones tiene que estar estrechamente ligados, por lo cual los valores de los componentes deben de permanecer en una pequeña tolerancia. Por esta razón la resistencia del oscilador debe ser especificada con una tolerancia del 1% y es usada para mantener exactitud y eliminar ajustes. Un potenciómetro ofrece otra posibilidad para mantener ajustado el mensaje.

La especificaciones del TMS3477 describen el tiempo de muestreo como una función de las resistencias de instalación y de la constante de tiempo externa RC Fig.4.3.5. Adicionalmente el TMS3477 puede ser manejado por una señal de reloj externa o cristal Fig.4.3.5.1. y Fig.4.3.5.2. Este circuito usa una estrecha tolerancia en sus componentes para operar con una frecuencia de muestreo de 32 Bits/seg.

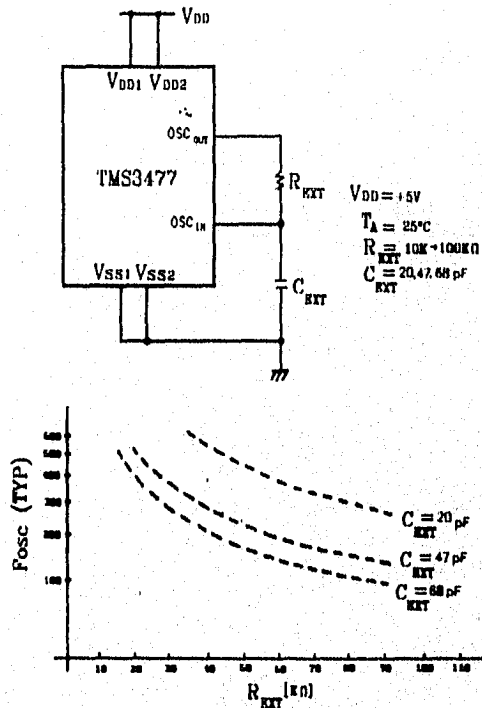


Fig.4.3.5. Oscilador R/C.

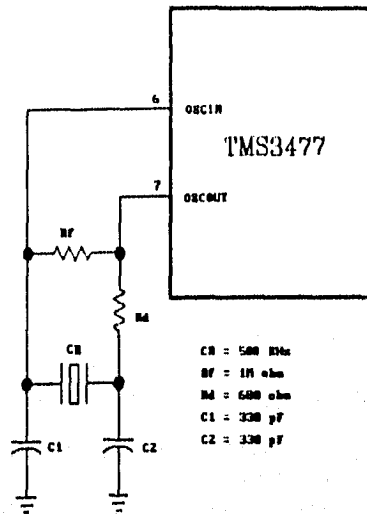


Fig.4.3.5.1. Oscilador con cristal.

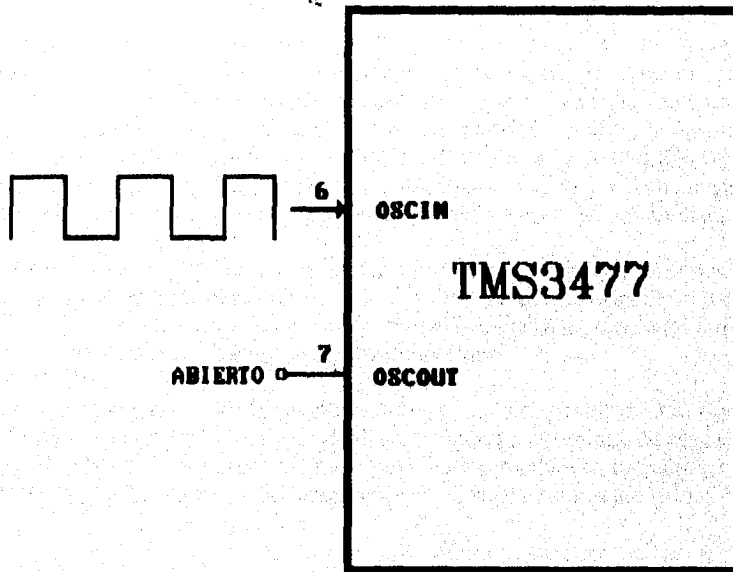


Fig.4.3.5.2. Reloj de entrada externa.

CONCLUSIONES.

Es factible que hoy en día existan personas que no llegan a comprender lo que significan los avances tecnológicos que se están dando en todo el mundo. Y lo que es peor un sector mas grande de la población mundial no tiene acceso a ellos. Me refiero a logros tan importantes como : el que una persona se pueda comunicar con otra persona en el otro extremo del mundo a través de un teléfono inalámbrico; el que un turista visite un país donde no se hable su idioma y a través de un traductor de bolsillo se pueda comunicar con las personas; o que un empresario ya no utilice los servicios de una secretaria ya que su computadora personal cuenta con un sistema en el que solo se tiene que dictar y la computadora escribe e imprime los documentos.

Todos estos ejemplos han sido posibles gracias a un puñado de personas interesadas en el progreso a través de la electrónica. Y para ello han tenido que realizar grandes esfuerzos para lograr su objetivos. Hoy en día están de moda los sistemas que pueden sintetizar la voz y aplicarlos en operaciones de la vida diaria (como los ejemplos anteriores). Esta ocasión tocamos el tema de un digitalizador que a bajo costo transforma la voz en un banco de datos formado por 2 simples valores que son el 1 y el 0. Y de esta manera dejar sorprendidos a propios y extraños que un simple circuito integrado CI realice tantas operaciones. Se presentaron los fundamentos teóricos de la transmisión digital de la señal de voz, por lo que la referencia a esta tesis servirá para entender los principios de funcionamiento de la tecnología en esta área de las telecomunicaciones, la cual empieza a tener auge en nuestro país, principalmente por la instalación de centrales telefónicas digitales.

El sistema diseñado y construido ilustró el proceso detallado de la conversión analógica-digital y la conversión digital-analógica. La circuitería se implementó con elementos del mercado nacional, y por su enfoque teórico, la estructura funcional del sistema es de suma utilidad en la digitización de señales analógicas que posean otras variantes a las consideraciones aquí mostradas, dado su diseño a bloques y haber incluido elementos de diseño adaptables a cada situación. El sistema de digitalización de voz desarrollado aquí puede ser usado como equipo didáctico en laboratorios de comunicaciones y electrónica, o como equipo de instrumentación en los requerimientos de campo.

Se señalaron las bases para realizar la interfase a otros dispositivos , por lo que una posible continuación al trabajo desarrollado en esta tesis se referiría a la realización de un reconocedor de voz (verificación e identificación de personas), un sistema inteligente que realice operaciones a través de la voz como en la educación o la ayuda a impedidos.

Por último quiero hacer mención de que en nuestro país no ha sido posible la utilización provechosa de todos estos grandes logros tecnológicos debido a la burocracia y sistemas de educación actuales. Por lo tanto se concluye que es buen momento para realizar cambios en los métodos educativos y dar oportunidad a quien realmente la merece y no quedar en el atraso tecnológico.

BIBLIOGRAFÍA

Information Transmission, Modulation and Noise

M. Schwartz

Mc Graw Hill, 3a. Ed., 1981.

Introducción a la Teoría y Sistemas de Comunicación

B. P. Lathi

Limusa, 1a. Ed. 1978.

Principles of Quantization

A. Gersho

IEEE Transactions on Circuits and Systems, Vol. CAS-25

Núm. 7, Jul. 1978. pp. 427-436.

Future Developments in Telecommunications

J. Martin

Prentice Hall, 2a. Ed., 1977.

Waveform Encoding of Voiceband Data Signal

J. B. O'Neal Jr.

Proceedings of the IEEE

Vol. 68, Núm. 2, Feb. 1980.

Differential PCM for Speech and Data Signals

J. B. O'Neal Jr. y R. W. Stroh

IEEE Transactions on Communications

Vol COM-20, Núm.5, Oct. 1972.

Delta Modulation Techniques and Devices

T. N. Saadawi

Innovations in Telecommunications

Part A, Academic Press, 1982, pp. 173-213.

Digital Telephony

J. Bellamy

John Wiley & Sons, 1982.

Digital Communications

K. Feher

Prentice Hall Inc., 1981.

A Comparison of Modulation Techniques for Digital Radio
J. D. Oetting,
IEEE Transactions on Communications, Vol. COM-27
Núm. 12, Dic. 1979, pp. 1752-1762.

Digital to Analog Conversion Techniques
D. F. Hoeschele Jr.
John Wiley & Sons, 1968.

Introduction to Operational Amplifier Theory and Applications
L. P. Huelsman y J. V. Wait
Mc Graw Hill, 1975.

Applications of Digital Signal Processing
A.V. Oppenheim
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA 1978

Automatic Speech Synthesis
D. O'Shaughnessy
IEEE Communications Magazine
Diciembre 1983

Current Perspectives in Digital Speech
R.E. Crochiere y J.L. Flanagan
IEEE Communications Magazine
Enero 1983

Digital Filters and the Fast Fourier Transform
B. Liu
Princeton University
Stroudsburg, Pennsylvania, USA

Digital Processing of Speech Signals
L.R. Rabiner y R.W. Schafer
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA 1978

Digital Signal Processing
L.R. Rabiner y Ch.M. Rader
IEEE Press
New York, N.Y., USA 1972

Electronic Speech Synthesis: Techniques, Technology and Applications
G. Bristow
M.C Weav Hill Book, Co.
R.R. Donelley and Sons, Co.
New York, USA 1984

Multirate Digital Signal Processing
R.E. Crochiere y L.R. Rabiner
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA
The Philosophy of PCM
B.M. Oliver, J.R. Pierce y C.E. Shannon
Proc. IRE, Vol 36i No. 11
Noviembre 1948

Speech Analysis
R.W. Shafer y J.D. Markel
IEEE press
New York, USA 1979

Speech Synthesis
J.N. Holmes
M. & B. Monograph, Mills and Boon Limited
London, Great Britain 1972