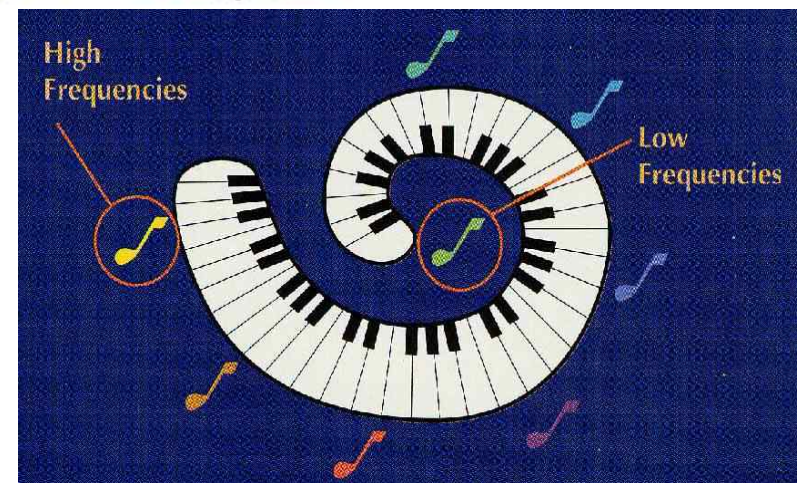
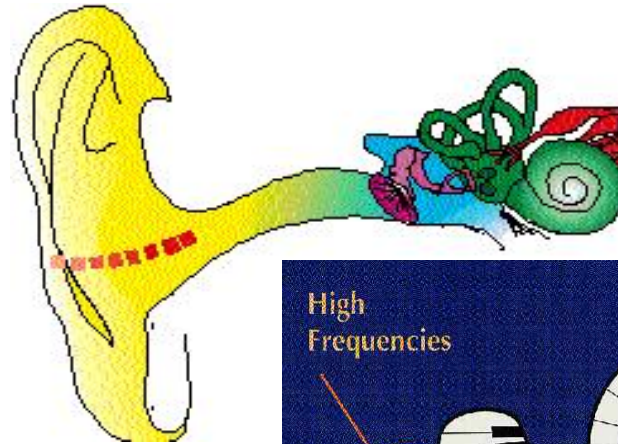
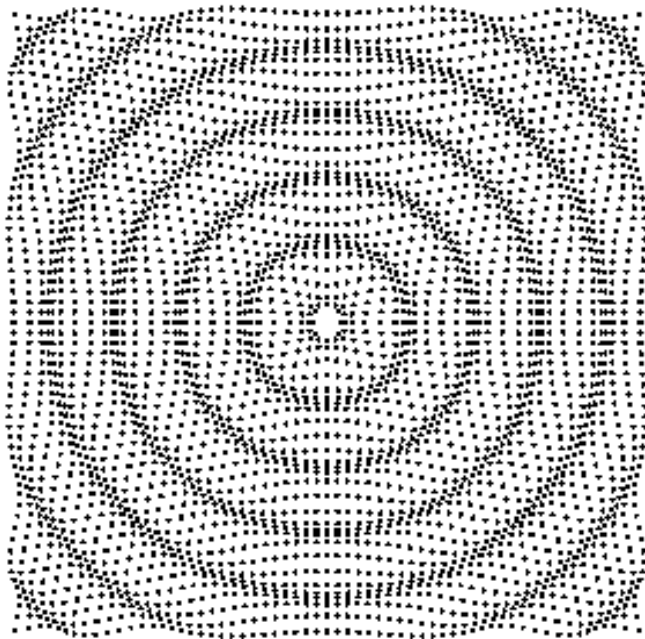
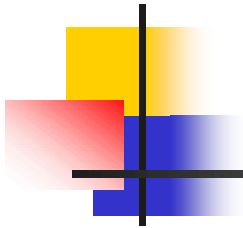


제9강. 음향 음성학의 기초 개념

수업 목표 :

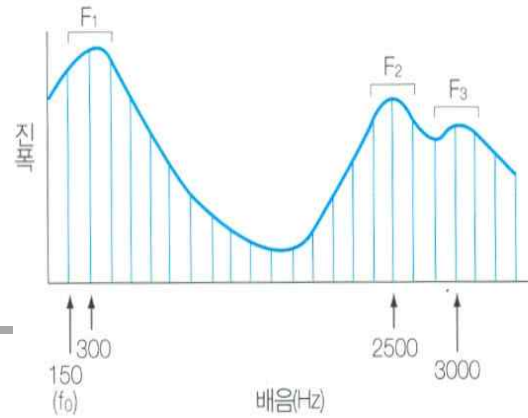
1. 음향음성학의 기본적인 개념과 용어를 이해한다.
2. 음향신호에 대한 3가지 서로 다른 표현 방법(파형, 스펙트럼, 스펙트로그램)에 대해 이해한다.
3. 디지털 녹음과 분석을 위한 기본적인 개념을 이해한다(표본화, 양자화, Nyquist 주파수).



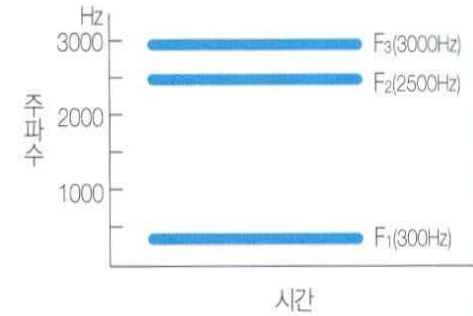


내용

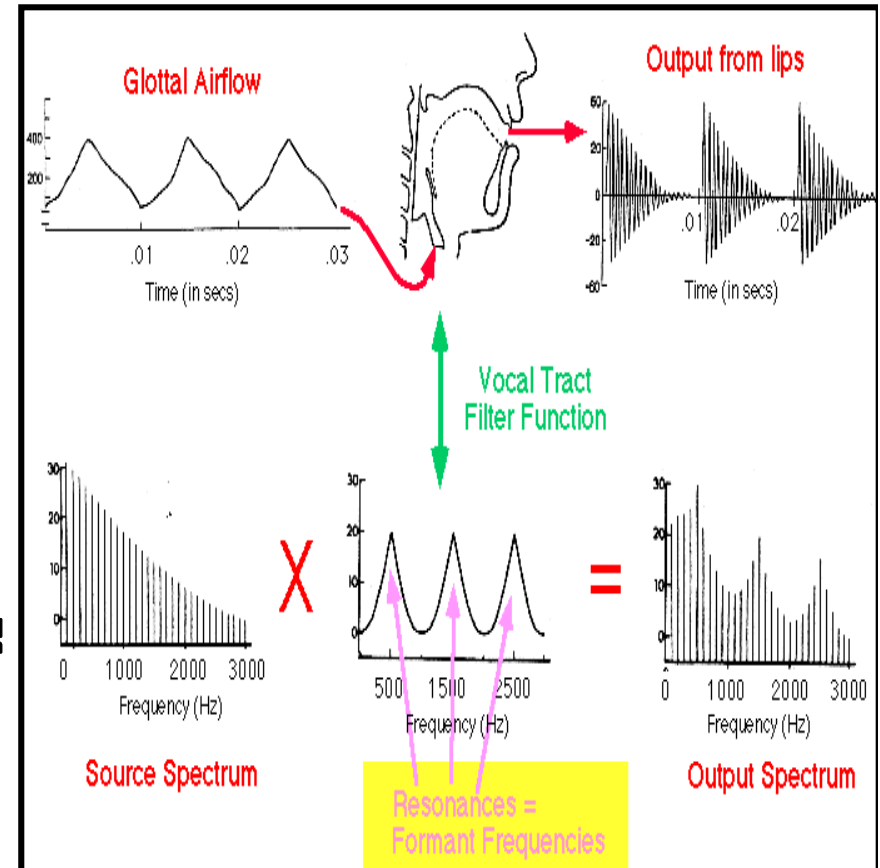
[i] 모음의 스펙트럼

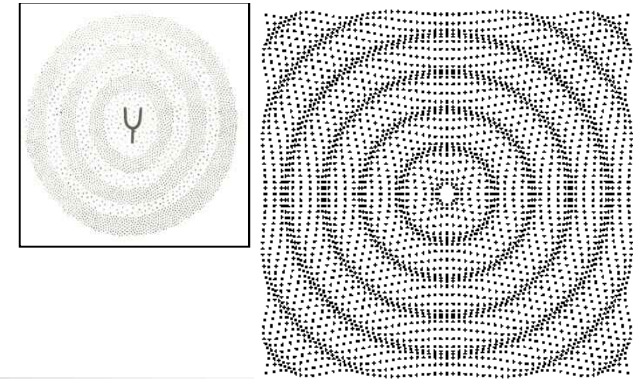


[i] 모음의 광역 스펙트로그램



1. 소리가 일으키는 파동 : 음파(sound wave, 파형(waveform))
2. 스펙트럼
3. 스펙트로그램
4. 음압, 강도, 데시벨(dB)
5. 디지털 녹음과 분석
 1. 디지털과 아날로그
 2. 표본추출률
 3. 양자화
 4. 디지털 신호의 스펙트럼, 스펙트로그램 분석

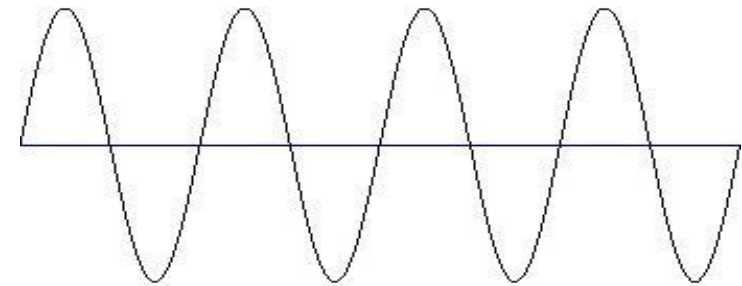




1. 소리가 일으키는 파동 : 음파

■ 단순파(simple wave)

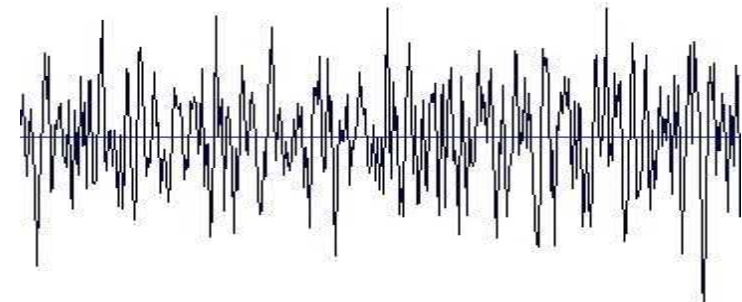
- 음파의 가장 단순한 형태
- 소리 굽쇠의 진동으로 만들어짐
 - 주기파 : 일정한 패턴이 반복되는 파형
 - X축 : 시간
 - Y축 : 진폭
 - $f = 1/T$



(순음)

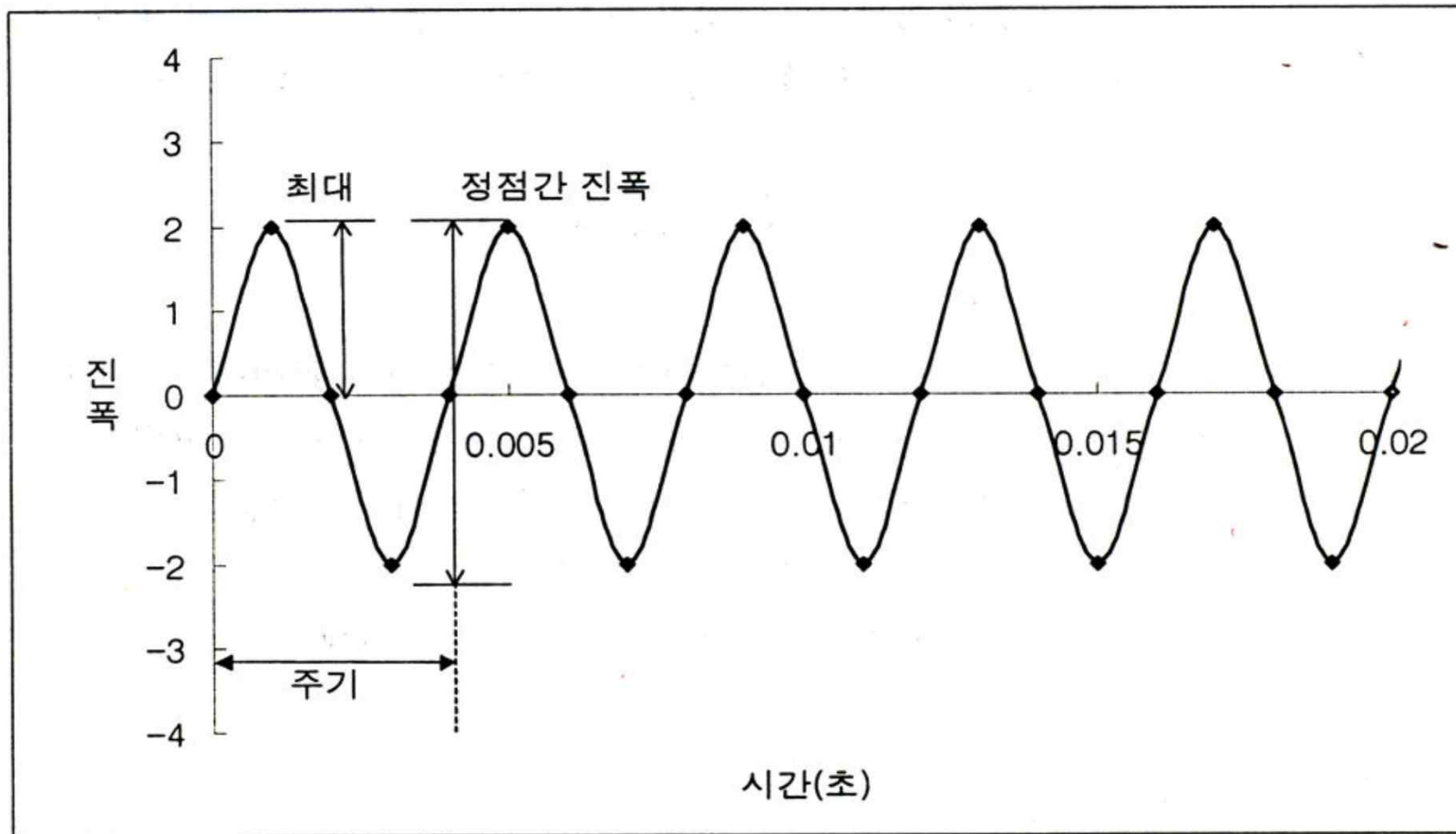
■ 복합파(complex wave)

- 복잡한 형태를 가진 파형
- 둘 이상의 요소가 복합된 형태의 파형
- 푸리에(**Fourier**), 18세기 프랑스 수학자 : 모든 주기파는 일정수의 단순파들의 합



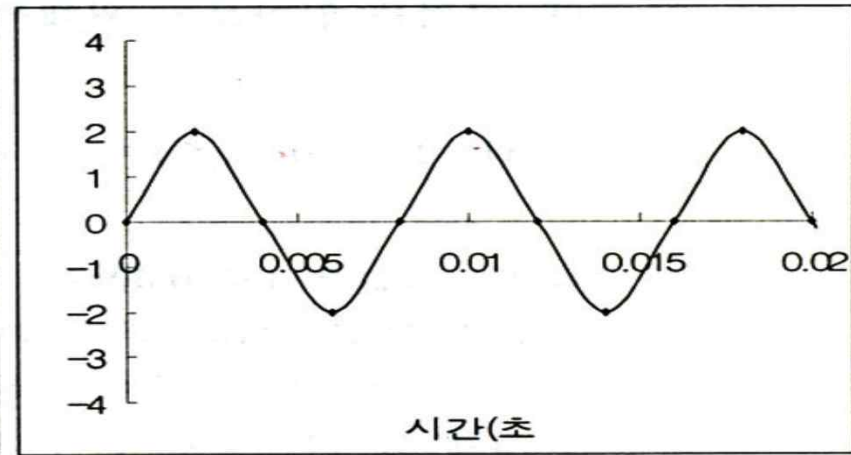
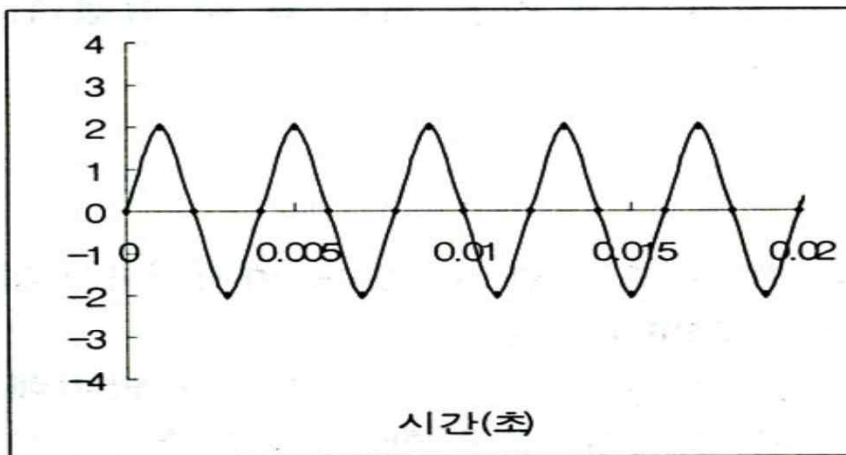
(복합음)

1.1 단순파의 예

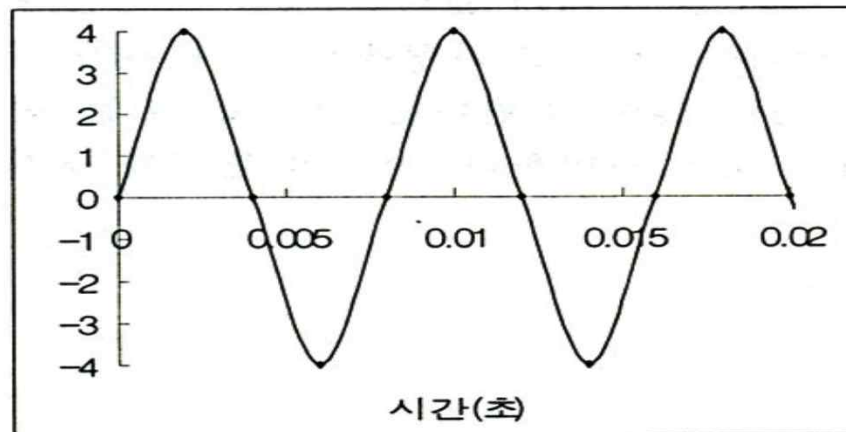


[그림 7-1] 단순파의 예

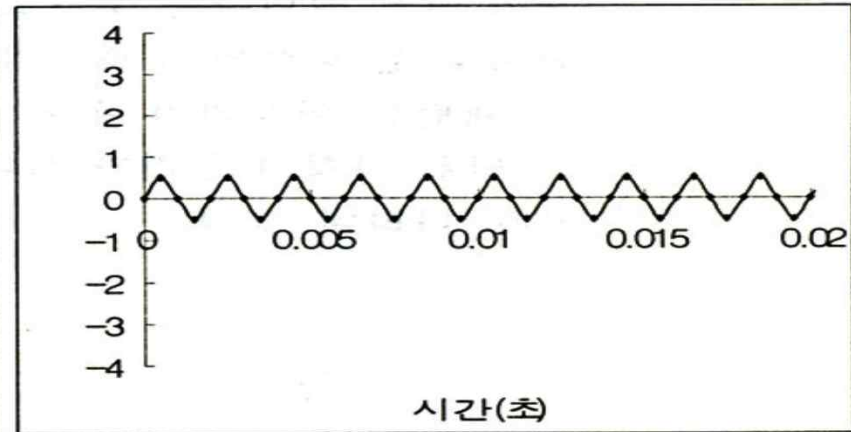
1.2 주파수, 진폭을 달리하는 다양한 종류의 파형들



ㄷ)

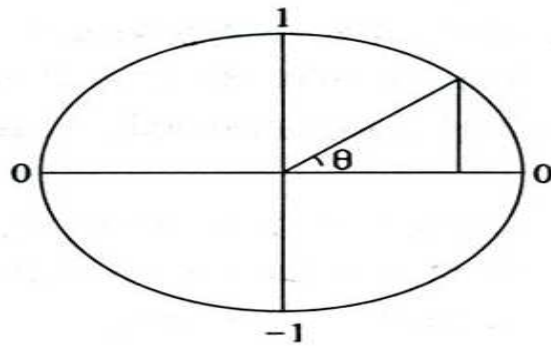


ㄹ)



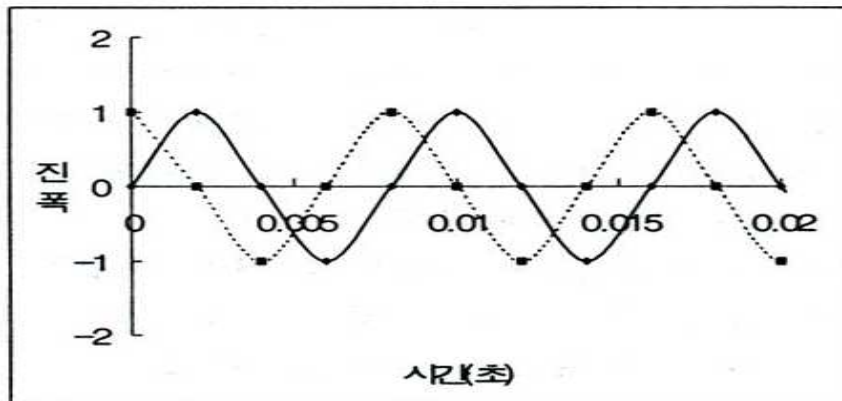
1.3 서로 다른 위상차를 보이는 사인 곡선들

ㄱ)

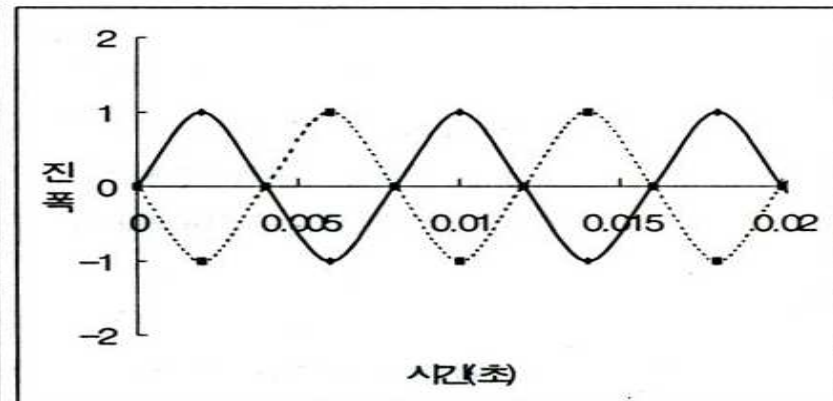


$\sin \theta$	
0°	0
45°	0.71
90°	1
135°	0.71
180°	0
225°	-0.71
270°	-1
315°	-0.71
360°	0

ㄴ)



ㄷ)



ㄱ)은 사인값을, ㄴ)은 90° 의 위상차를 갖는 경우를, ㄷ)은 180° 의 위상차를 갖는 경우를 각각 보인 것이다.

1. SBS 놀라운대회 스타킹 : 유리잔으로 내는 정말 고운소리

<http://netv.sbs.co.kr/newbox/newbox.jsp?uccid=10000341248&boxid=>

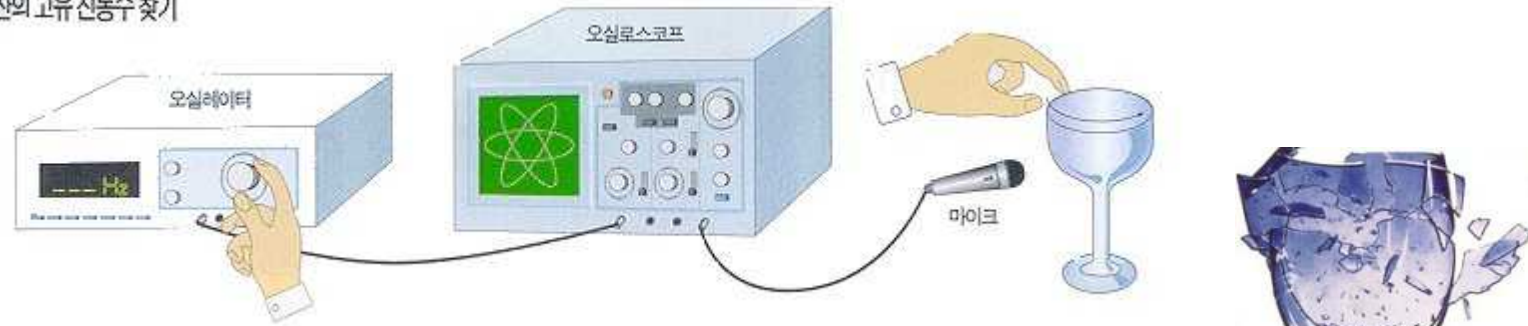
2. 스펀지(김종서 유리잔 깨는 실험)

3. 공명의 위력 실험(다리붕괴)



1.3.1 소리로 유리컵 깨뜨리기

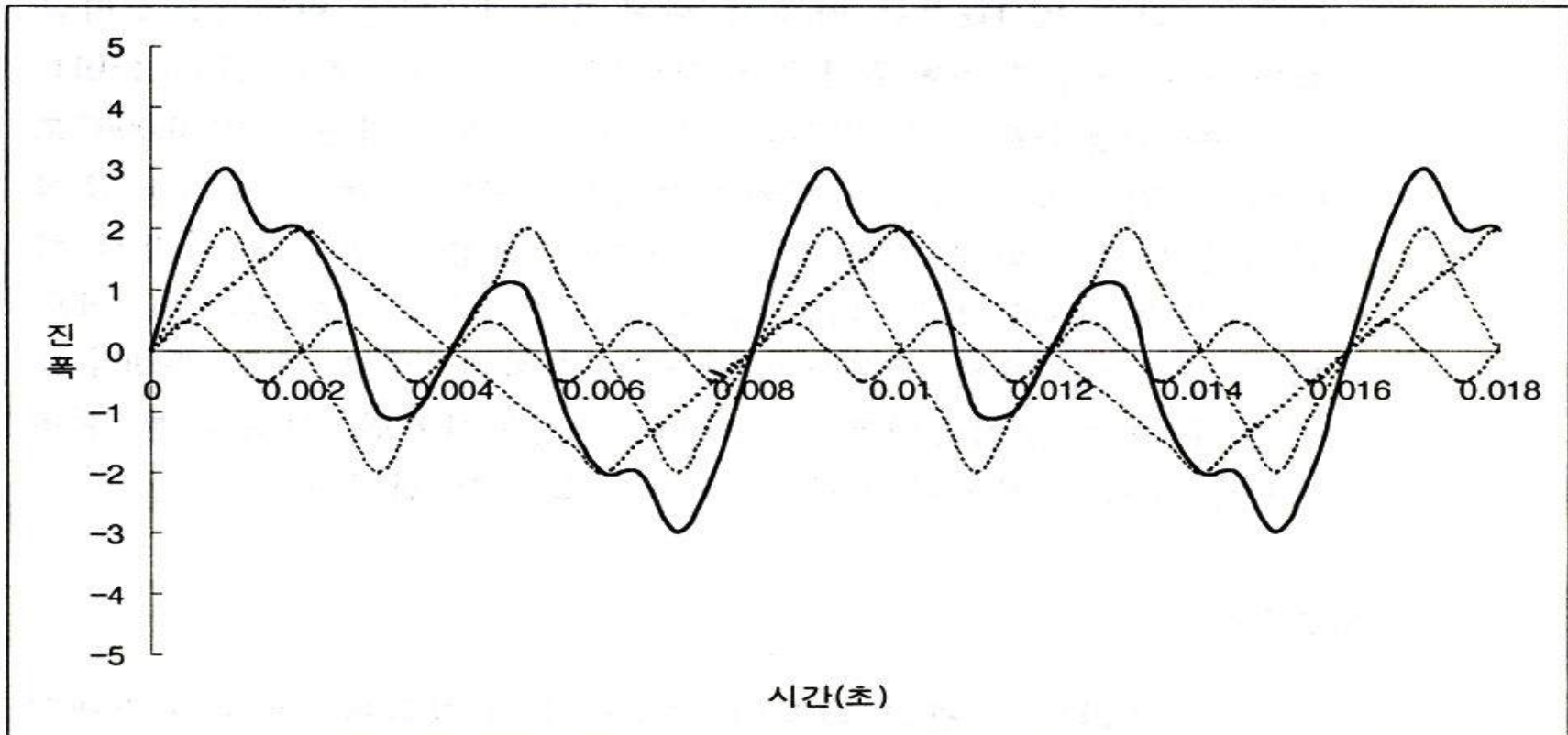
유리잔의 고유진동수 찾기



오실레이터로 유리잔의 고유진동수와 같은 진동수를 발생시켜 컵을 깨뜨릴 수 있다.



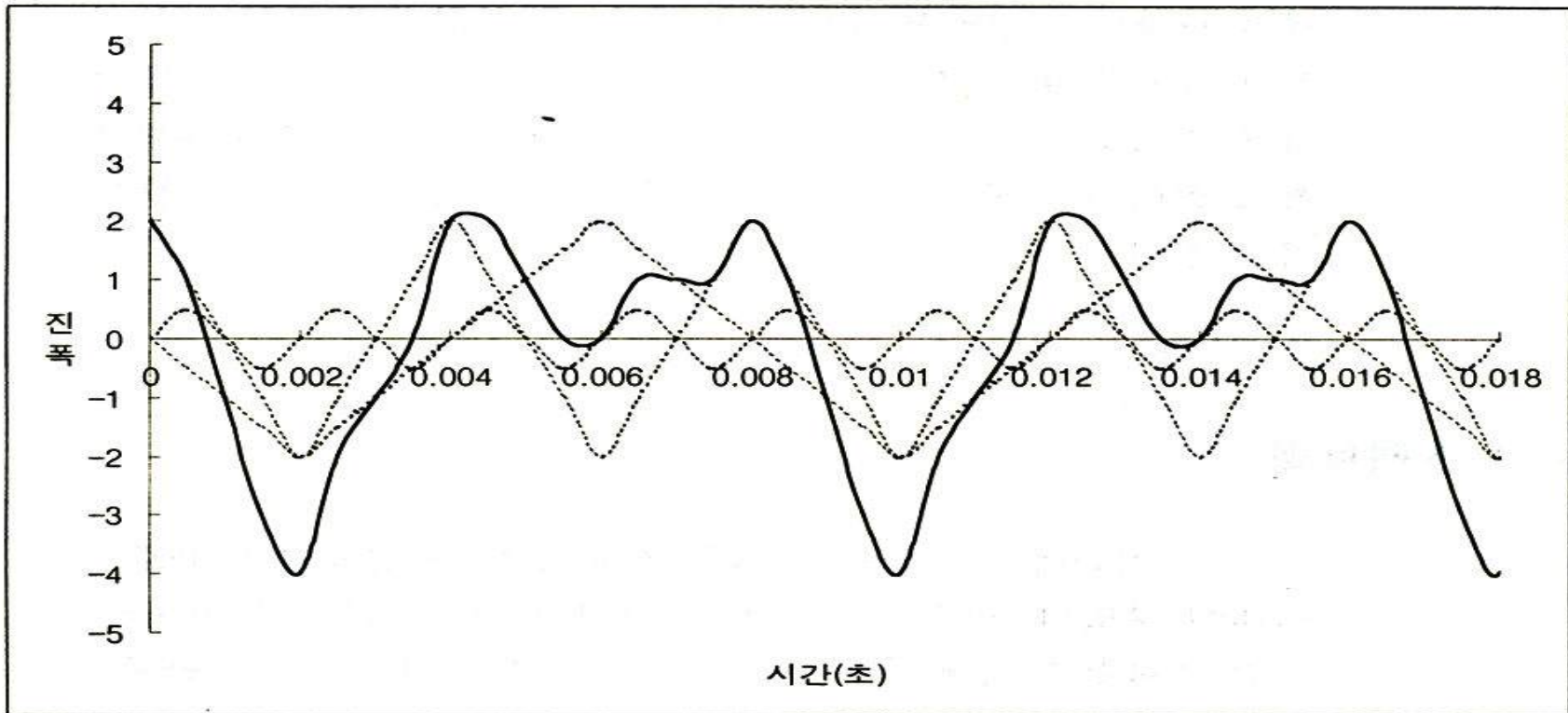
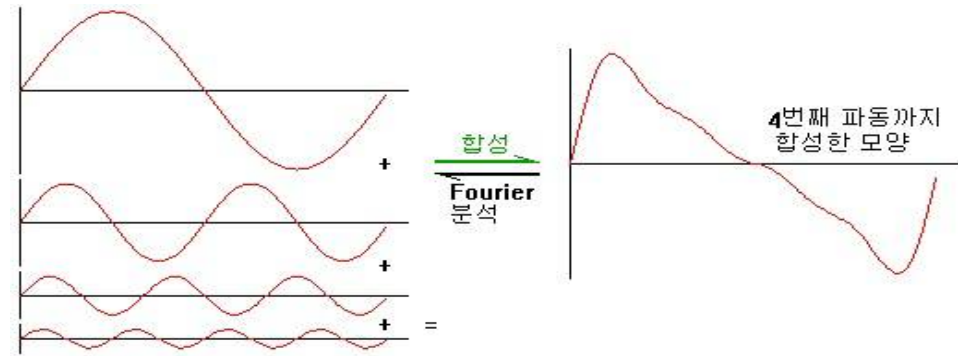
1.4 복합파의 예



위 그래프에서 점선은 단순파를, 실선은 복합파를 각각 나타낸다.

[그림 7-4] 그림 7-2의 ㄱ), ㄴ), ㄹ)에 보인 세 단순파의 복합으로 만들어진 복합파의 예

1.5 복합파의 예



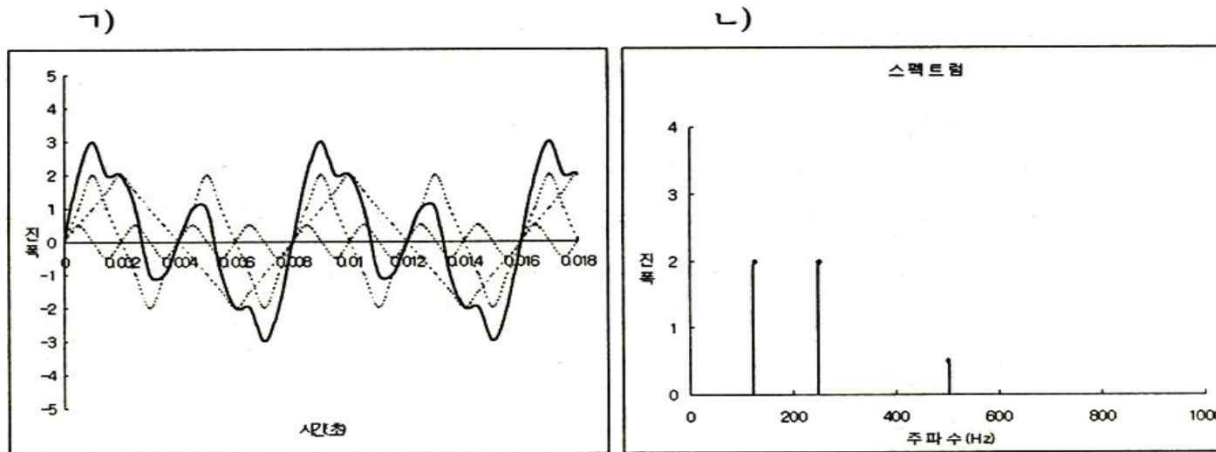
위 그래프에서 점선은 단순파를, 실선은 복합파를 각각 나타낸다.

[그림 7-5] 그림 7-2의 ㄱ), ㄴ), ㄷ)에 보인 세 단순파와 각각 90° , 180° , 0° 의 위상차를 가진 세 단순파의 복합으로 만들어진 복합파의 예

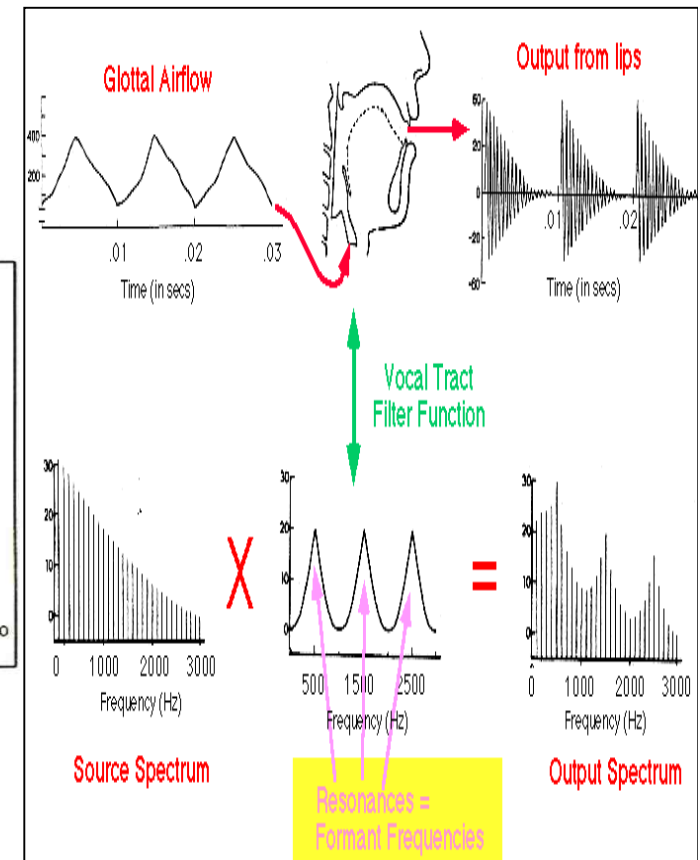


2. 스펙트럼(spectrum)

- 스펙트럼 : 음향신호를 주파수, 진폭으로 분석하여 보여줌
 - 진폭, 주파수 도면으로써 복합파를 분석하여 그 복합파가 어떠한 단순파들의 합으로 이루어진 것인가를 보여주는 방법
- 요리분석기가 분석한 냉커피의 스펙트럼

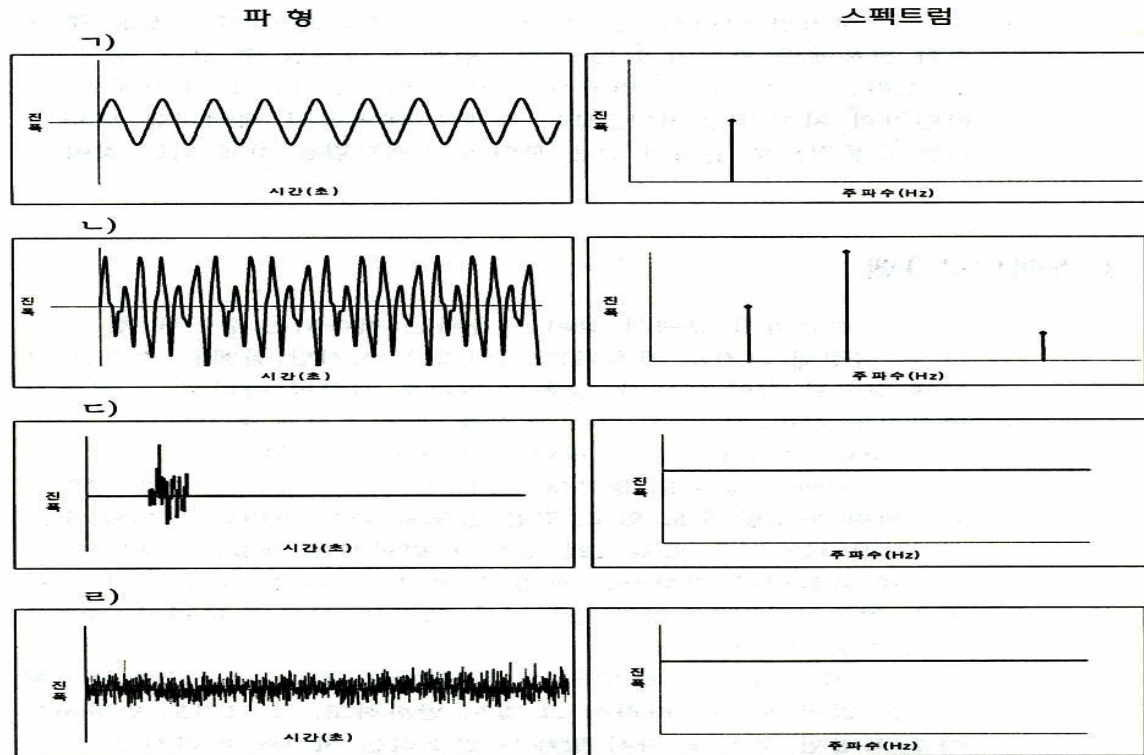
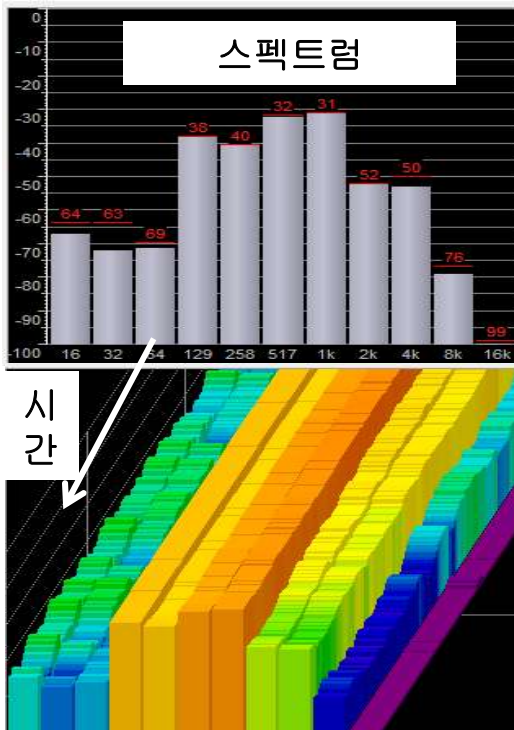


[그림 7-7] ㄱ)은 복합파의 파형을, ㄴ)은 이 복합파의 스펙트럼을 보여 준다.



2.1 여러 가지의 스펙트럼

- 요리분석기가 분석한 냉커피의 스펙트럼
 - 선스펙트럼
 - 연속스펙트럼

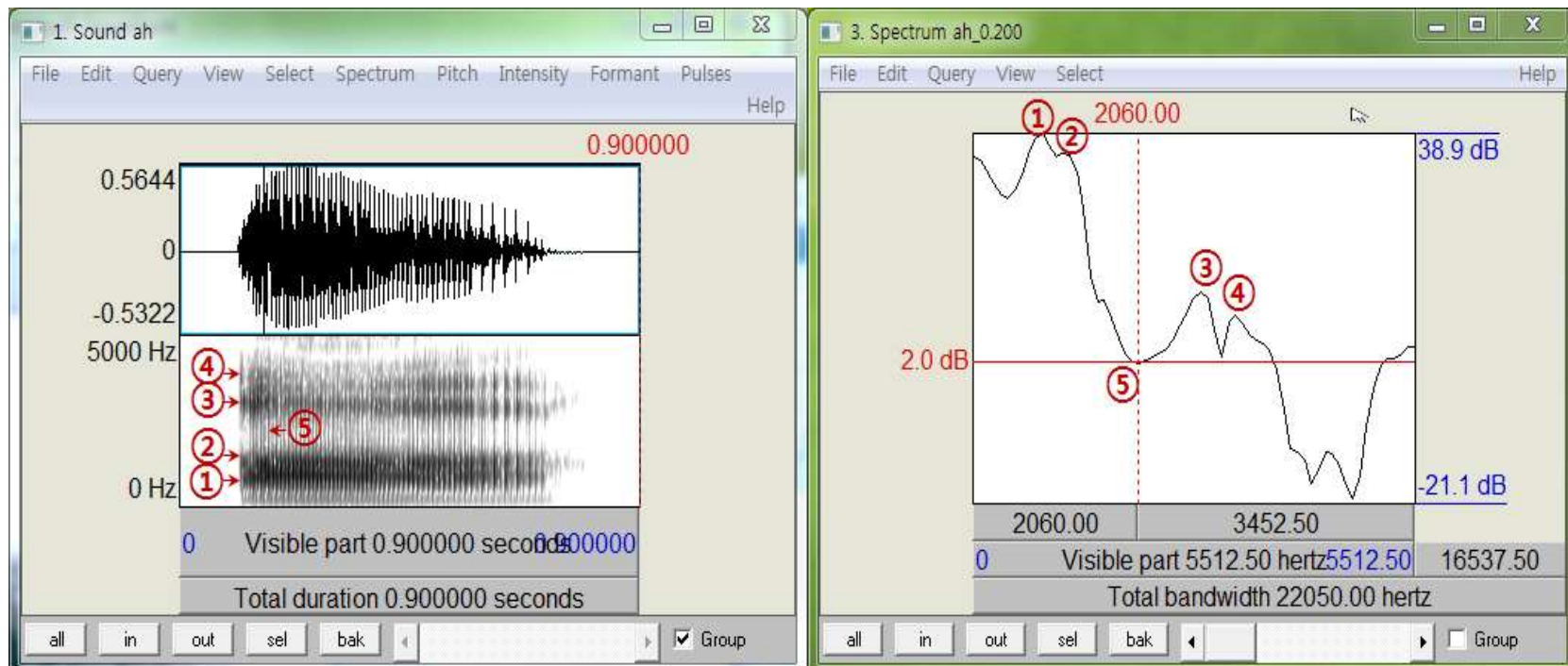


ㄱ), ㄴ)은 주기파의 파형과 스펙트럼을 보여 주고 있는데, 여기서 ㄱ)은 단순파의 경우를, ㄴ)은 복합파의 경우를 보인 것이다. ㄷ), ㄹ)은 비주기파의 파형과 스펙트럼을 보여 주고 있는데, 여기서 ㄷ)은 순간음의 경우를, ㄹ)은 소음의 경우를 보인 것이다.

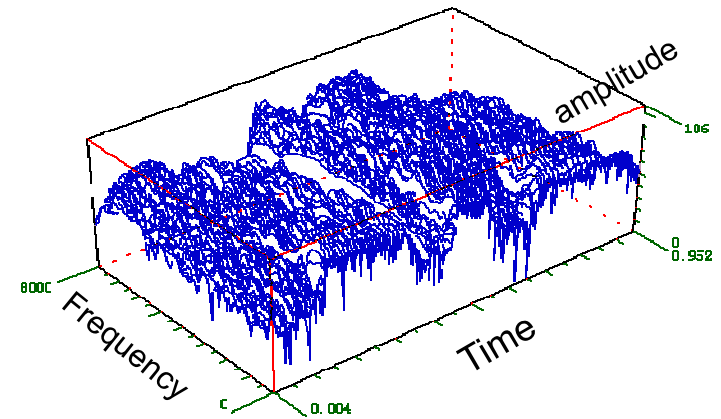
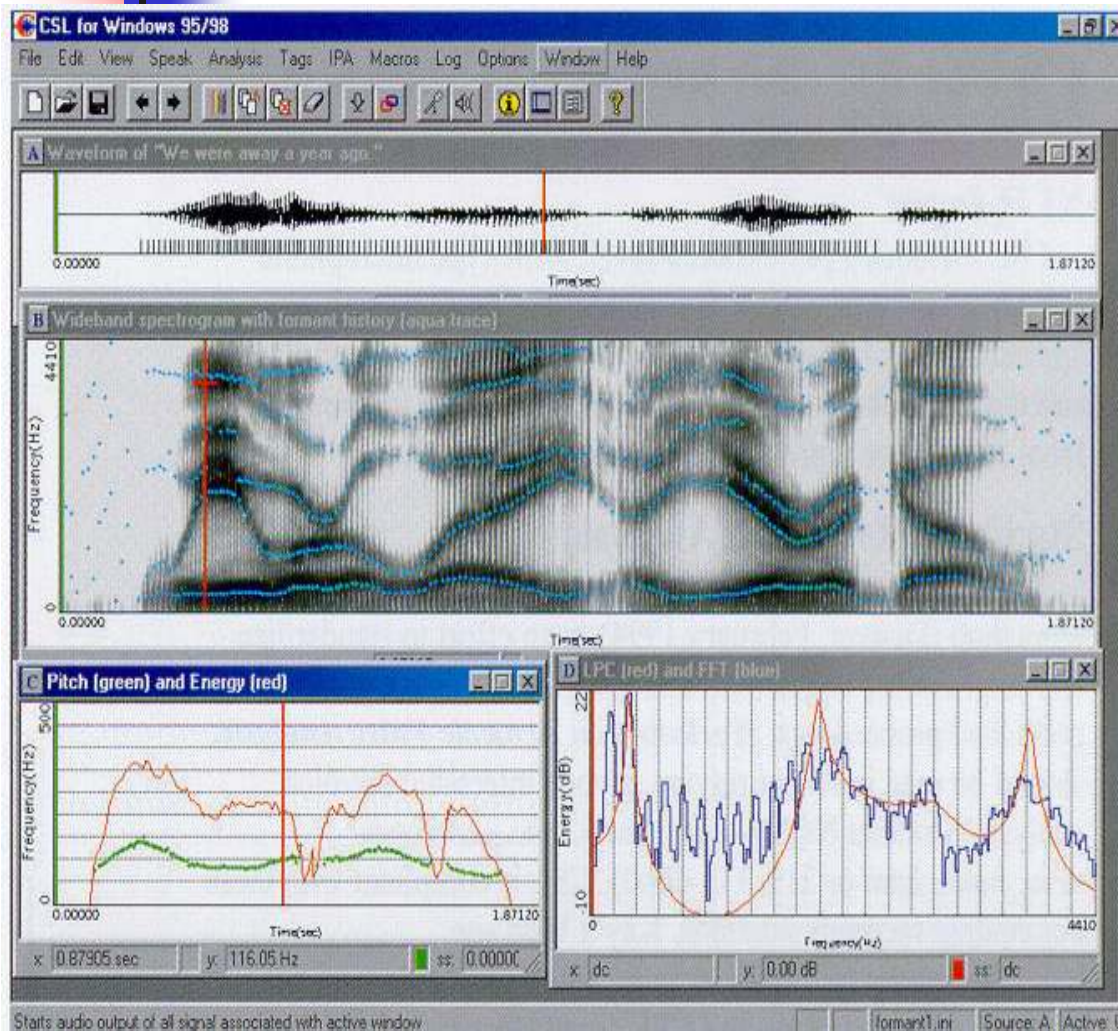
[그림 7-8] 여러 가지 음파의 스펙트럼

3. 스펙트로그램(spectrogram; SPG)

- 스펙트로그램 : 음향 신호를 주파수, 진폭(강도), 시간으로 분석하여 얻어진 그림
 - 주파수와 진폭의 시간에 따른 변화를 보여주는 삼차원적인 그림
 - X축 : 시간
 - Y축 : 주파수
 - Z축 : 진폭

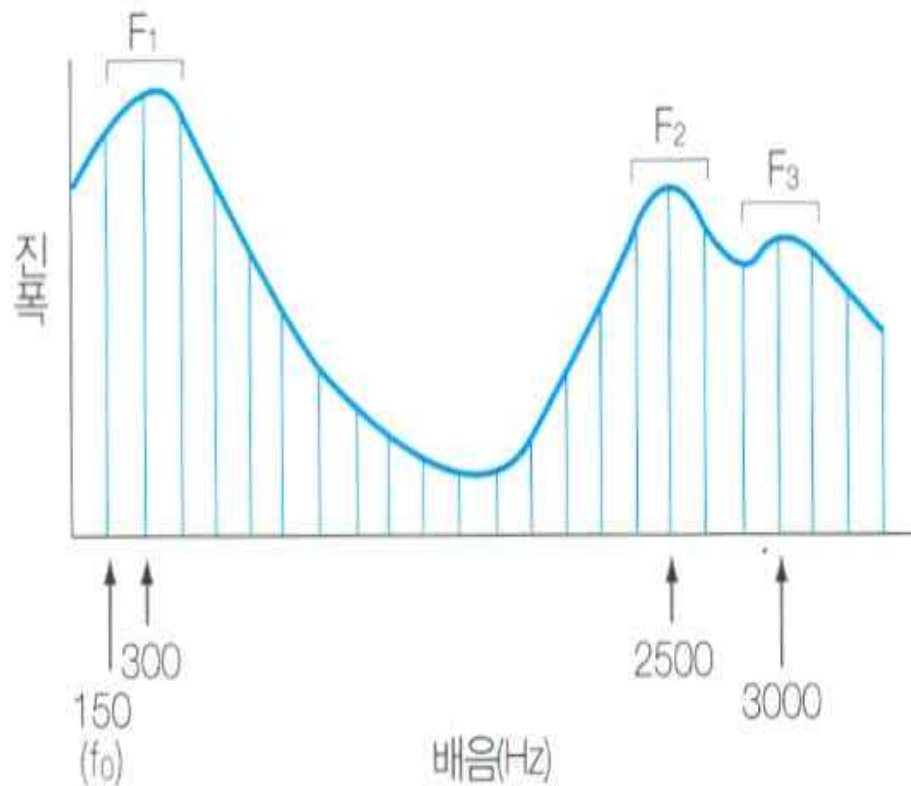


3.1 스펙트로그램(spectrogram; SPG)

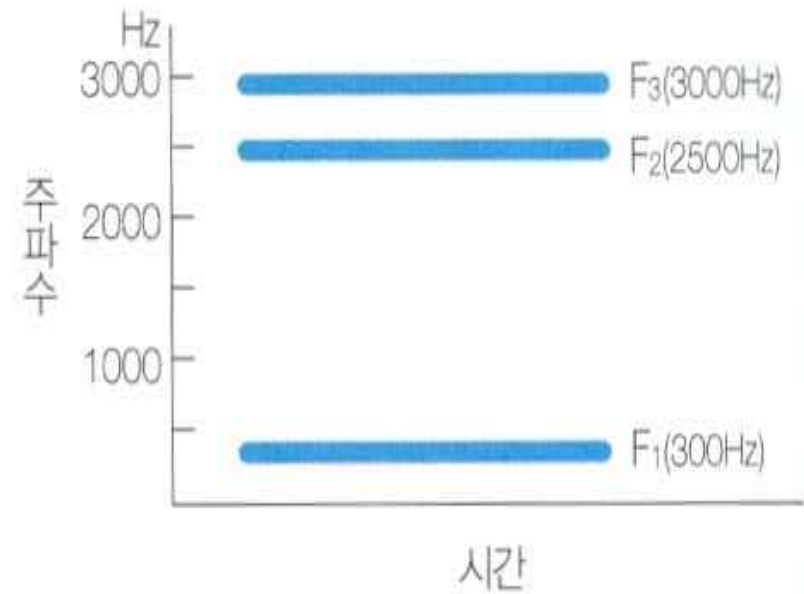


3.1.1 Spectrum vs Spectrogram

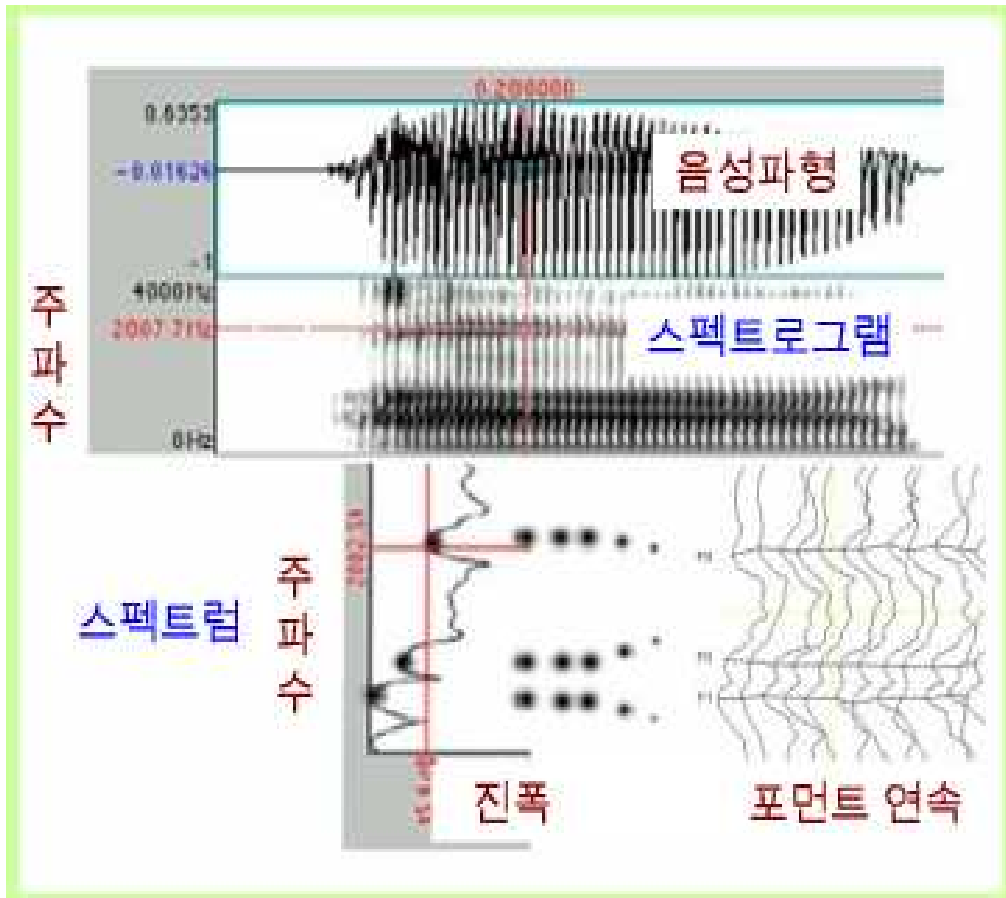
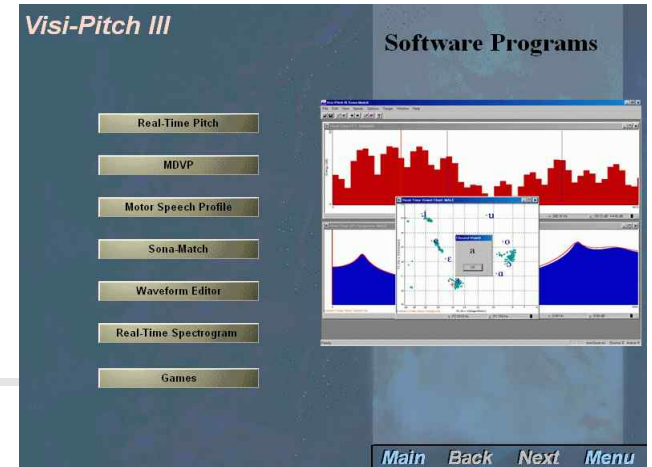
[i] 모음의 스펙트럼



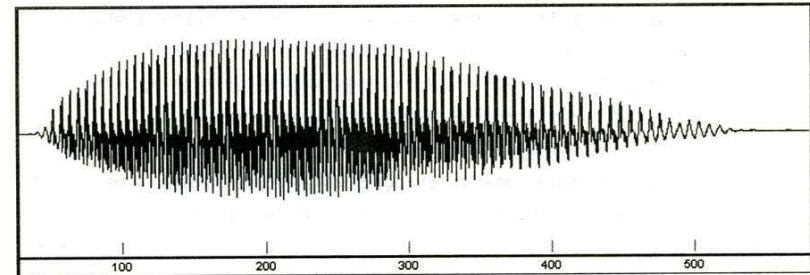
[i] 모음의 광역 스펙트로그램



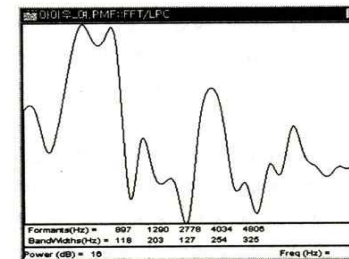
3.1.1.1 소리를 시각화하는 세 가지 방법



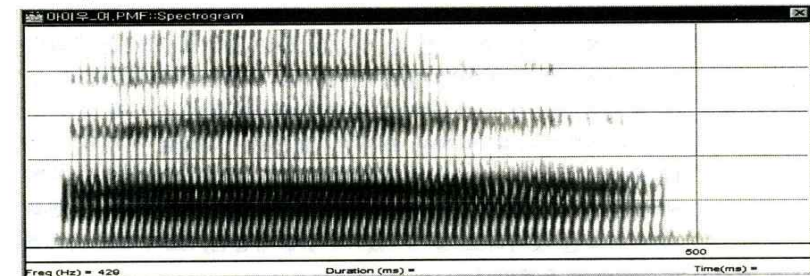
ㄱ) 음파



ㄴ) 스펙트럼



ㄷ) 스펙트로그램



[그림 7-9] 소리를 시각화하는 세 가지 방법

3.1.2 윈도우 길이와 주파수 대역폭

- **윈도우 길이(window length) :** 분석 대상 파형의 길이를 정하는 일 p.143
 - 분석 대상 파형의 길이를 비유적으로 표현하는 것
 - 창문의 너비가 넓으면 긴 길이의 좌우 풍경을 볼 수 있음 : [그림 7-12 ㄴ]
 - 창문의 너비가 좁으면 짧은 길이의 좌우 풍경을 볼 수 있음 : [그림 7-12 ㄷ]

- **주파수 대역폭(bandwidth) :** 복합파의 구성 성분의 종류와 양을 알아보기 위해 설정된 여과기의 폭
 - 윈도우 길이가 길면 주파수 대역폭이 좁고, 윈도우 길이가 짧으면 주파수 대역폭이 넓음
 - 윈도우 길이가 길고, 주파수 대역폭이 좁다는 것
 - 짧은 시간 내에 일어나는 변화에 대해서도 충분한 정보를 얻을 수 있어서 시간에 따른 주파수 구성 성분이 어떠한 변화를 겪는지에 대한 자세한 정보를 얻을 수 있음
 - 협역 스펙트로그램
 - 장점 - 파형을 구성하는 개별 성분 주파수가 무엇인지, 그 진폭은 어느 정도인지에 대한 정보를 얻을 수 있음
 - 단점 - 어느 주파수 대역에 에너지가 적은지에 대한 정보를 알기 어렵고, 짧은 시간 동안 벌어지는 성분 주파수의 종류와 양이 어떻게 변화하는지에 대한 정보를 얻기 어려움
 - 윈도우 길이가 짧고, 주파수 대역폭이 넓다는 것
 - 광역 스펙트로그램
 - 장점 - 어떤 주파수 대역에 에너지가 몰려 있는지에 대한 정보를 얻기 쉬워서 포먼트 분석에 유리함
 - 단점 - 파형을 구성하는 개개의 성분 주파수에 대한 정보를 얻기 어려움

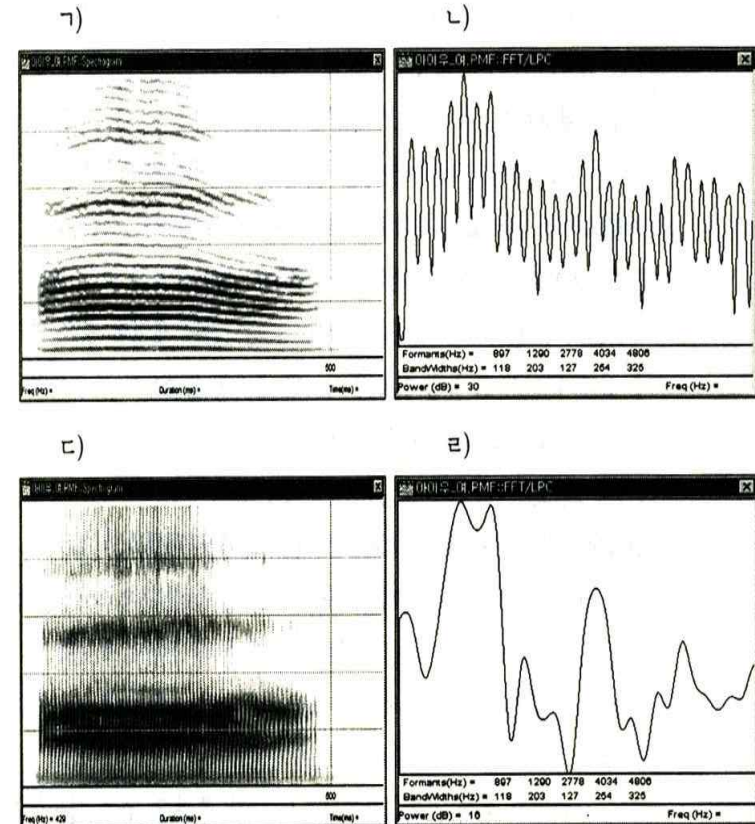
3.1.3 협역 스펙트로그램과 광역 스펙트로그램

- 협역 스펙트로그램(narrow-band spectrogram)

- 좁은 주파수 대역폭으로 음향신호를 분석
- 가로줄의 가는 선이 관찰
- 주파수 구성 성분의 변화 관찰
- 주파수 관련 정보를 얻기 용이

- 광역 스펙트로그램(wide-band spectrogram)

- 넓은 주파수 대역폭으로 음향신호를 분석
- 세로줄이 관찰 되고, 가로로 굵은 띠가 관찰
- 포만트(formant, 혹은 공명주파수) 관찰
- 시간 정보 얻기 용이



가)과 다)은 각각 모음 [a]의 협역, 광역 스펙트로그램을, 나)과 라)은 각각 가)과 다)의 스펙트럼 조각을 보여 준다.

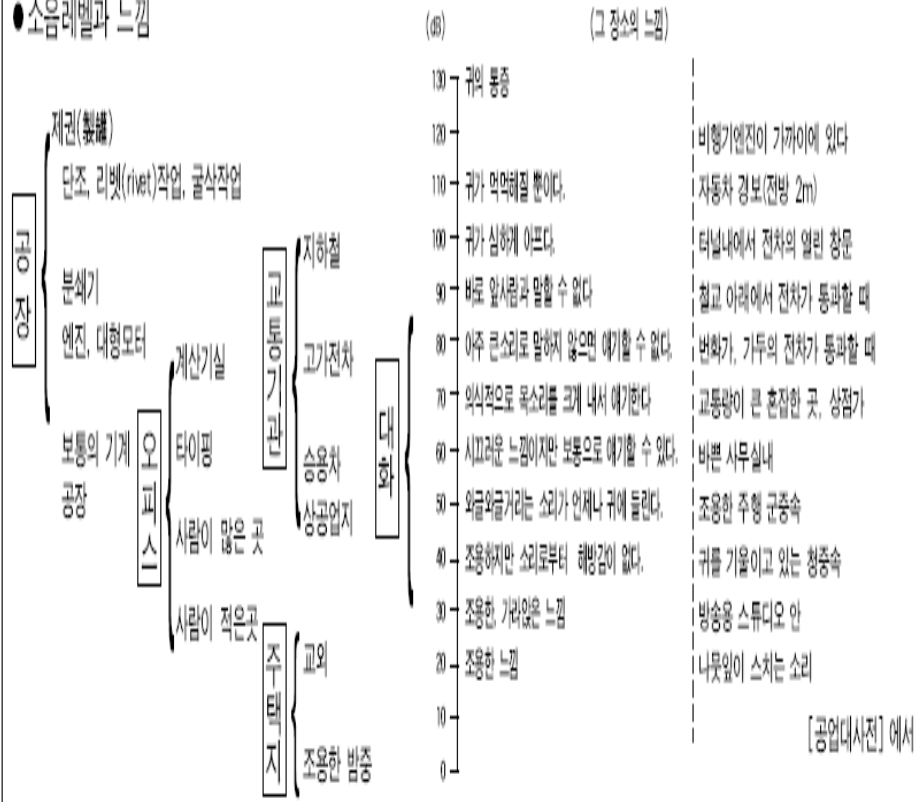
[그림 7-11] 협역 스펙트로그램과 광역 스펙트로그램

4. 음압, 강도, 데시벨(dB)

- 소리의 진폭은 음압 p.158
 - 음압 : 한 공기 분자가 진동에 의해 인접 공기 분자를 미는 힘
- 강도 : 파형이 전달되는 힘
- 단위 : dB
- SPL(sound pressure level)
 - 강도의 수준이 음압의 수준
- 현재 우리가 사용하는 강도 기준은 상대적인 강도 수준
 - 속삭이는 소리: 30dB SPL
 - 낮은 목소리의 대화: 40-50dB
 - 보통 발화: 60-70dB
 - 고탈 소리: 100dB

● 주변에서 빈번하게 접하는 소음의 개략값
 다음 그림은 주변에서 흔히 접하는 소음의 예이다. 이에 의해서 소음 수준의 개략을 알 수 있다.

● 소음레벨과 느낌



[공업대사전] 에서



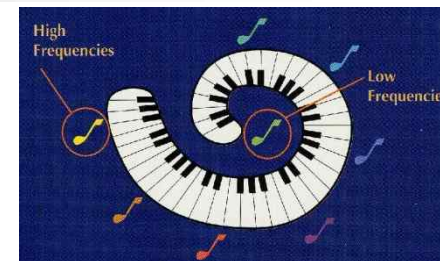
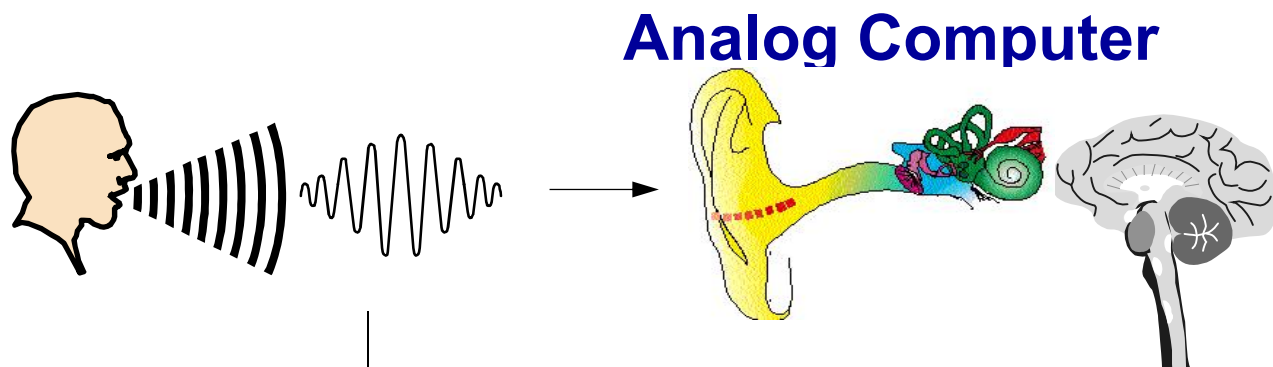
5. 디지털 녹음과 분석

- **디지털(digital)**
 - 숫자로 표기
 - 불연속적인 숫자 정보
 - **0**과 **1**로 표현되는 것

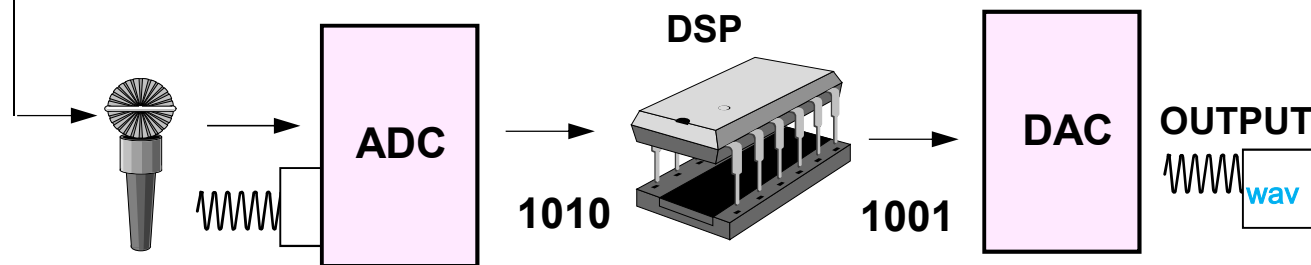
- **아날로그(analogue)**
 - 연속적인 숫자 정보
 - 실제 원음 원자료 그대로의 표현

5.1 디지털 신호처리(Digital Signal Processing; DSP)

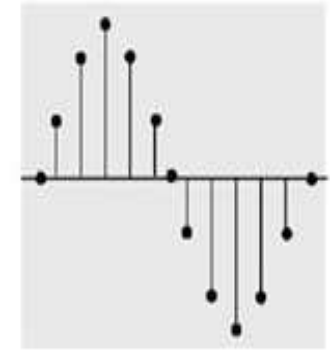
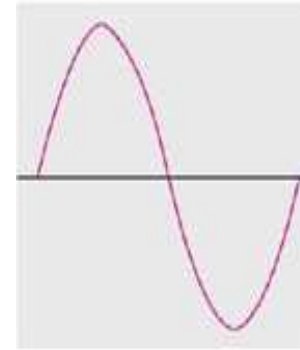
- 컴퓨터 음성녹음의 원리 -



Digital Computer

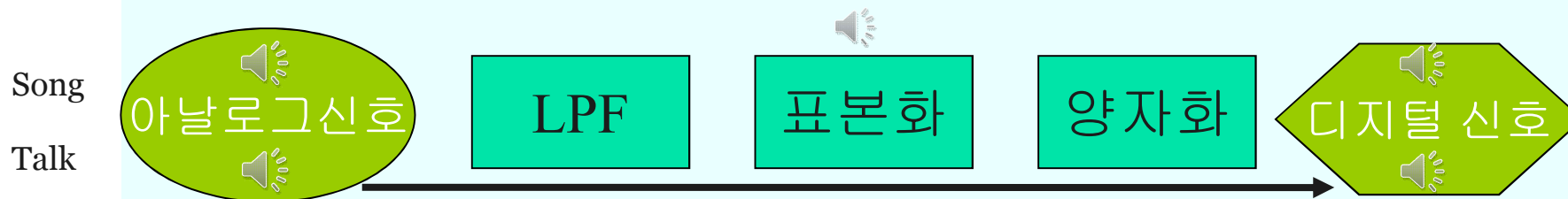


5.1.1 아날로그-디지털 변환 I



■ 신호의 표본화(sampling) 과정

- 표본화(sampling) : 아날로그 신호를 디지털 신호로 변환
- 신호의 표본화(sampling) 과정; 시간 축의 디지털화
- 아날로그 신호 정보를 0과 1로 표시되는 디지털 신호로 변환하여 수학적 연산에 의해 처리할 수 있도록 한 데이터 처리방법



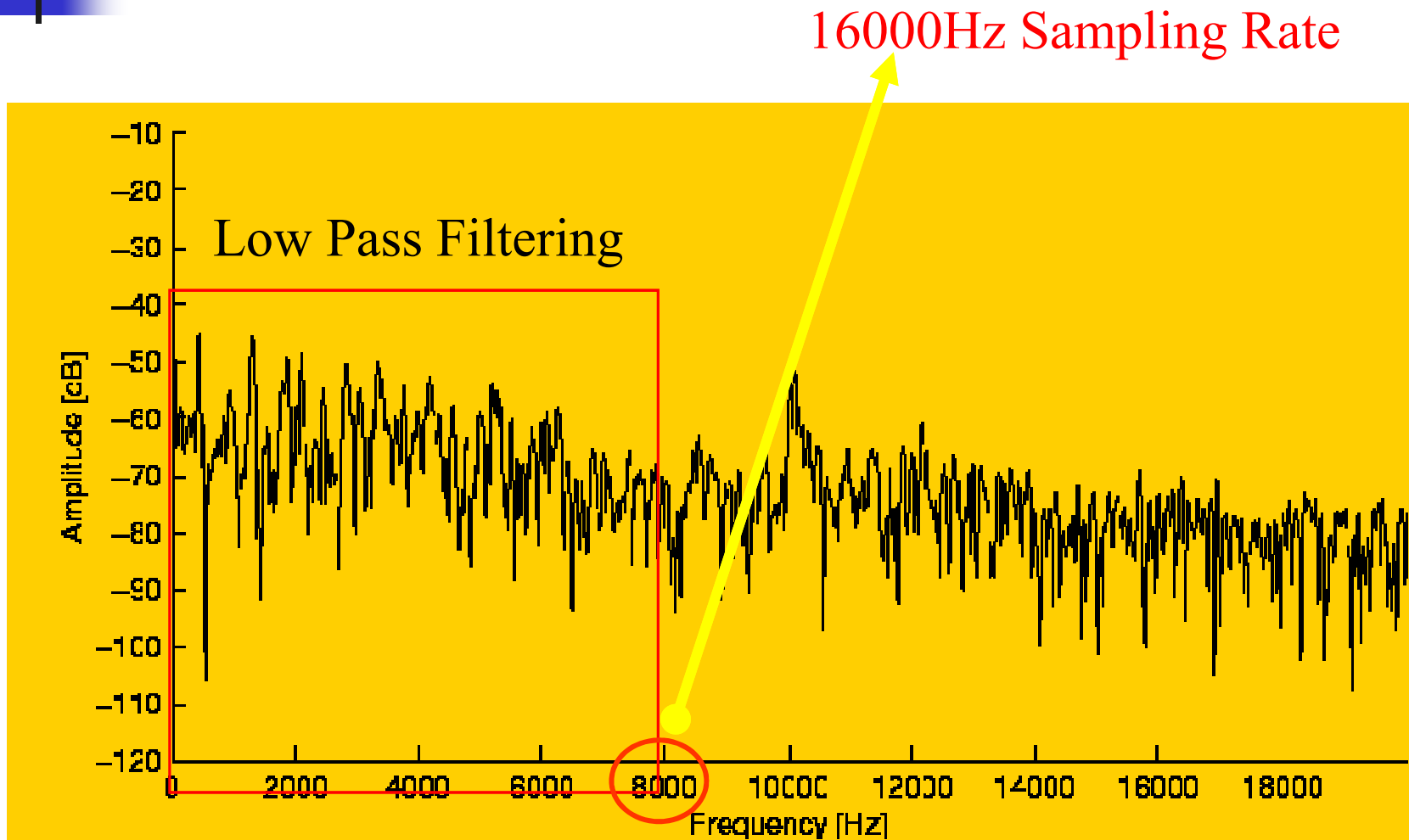
- LPF: 저역통과 필터
- 표본화: Sampling
- 양자화: Quantization

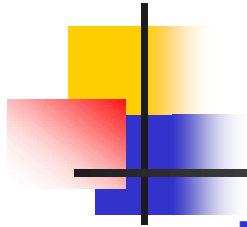


5.2 표본 추출률(sampling rate)

- 표본화(sampling)
 - 연속적인 원신호를 일 초당 몇 개의 점으로 저장할 것인가를 결정하는 것
- 표본추출률이 높다는 것은 그 만큼 점을 찍어서 원신호의 정보를 저장하는 것
- 파형의 시간축과 관계
 - 표본추출률의 적정선
 - 원하는 정보를 모두 얻으면서 저장용량은 최소가 되는 지점
- 최소한의 표본추출률은 원신호에서 우리가 관심을 두고자 하는 최대주파수 값의 2배가 되도록 함
- 나이퀴스트 주파수(Nyquist frequency) : 표본추출률의 절반 값

5.2.1 표본주파수의 선택(나이퀴스트 주파수)

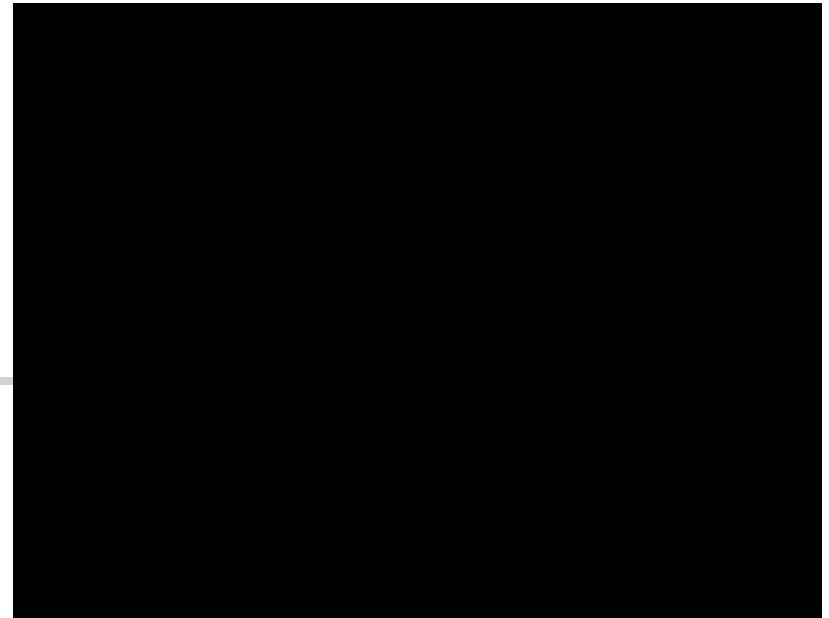


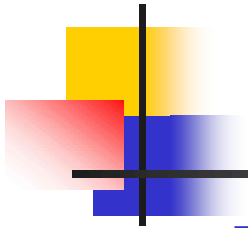


■ 표본률(sampling rate)

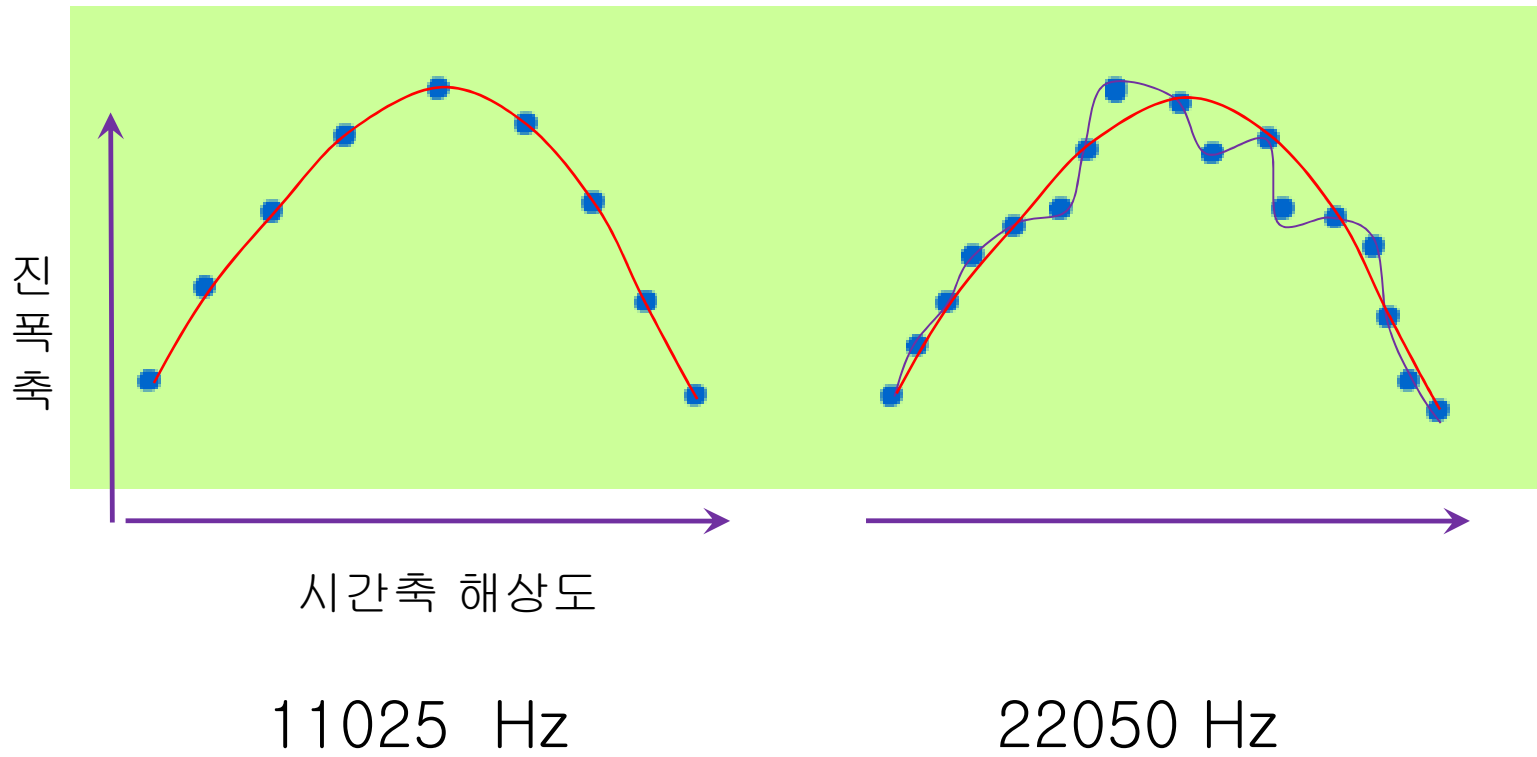
- Sample (사과상자 안에서 한 개 꺼냄)
- 1초를 몇 등분하여 소리의 진폭을 저장하는가?
 - 11025 Hz 는 1/11025초마다 하나씩 저장
 - 22050 Hz 는 1/22050초마다 하나씩 저장
- 나이퀴스트 주파수(Nyquist frequency) : 표본속도의 반에 해당하는 주파수 정보가 모두 저장됨
 - 사람의 청각범위: 20~20000 Hz
 - 44100 Hz 표본속도로 하면 22050 Hz까지 모든 정보를 저장

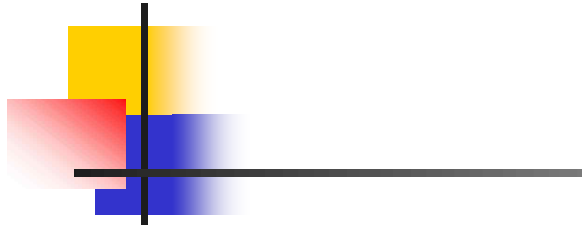
➡ 우리가 들을 수 있는 소리 모두 저장함





■ 표본 속도 비교

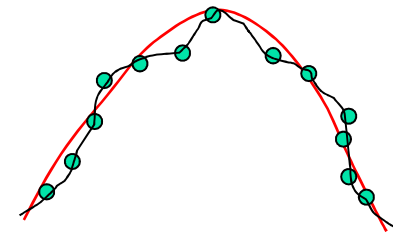
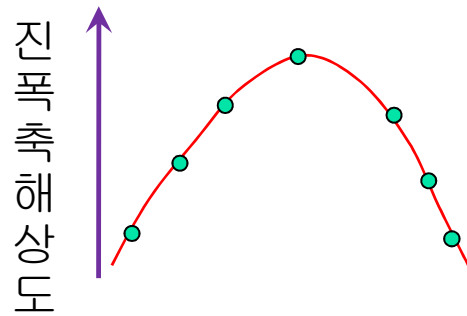




- 비트 수에 따른 진폭 해상도

- 8비트 $2^8=256$
- 16비트 $2^{16}=65536$
- 32비트 $2^{32}=4294967296$
- 64비트 $2^{64}=1.84467E+19$ (E+19는 10^{19} 를 나타냄)
= 1.84467×10^{19}

* 16비트로도 우리 청각기관에 필요한 충분한 정보임



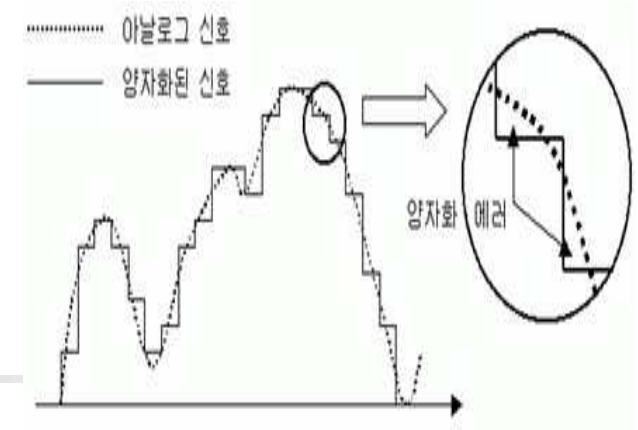


5.2.2 음성녹음: 기본적인 요구 조건

■ 기록장비와 매개체

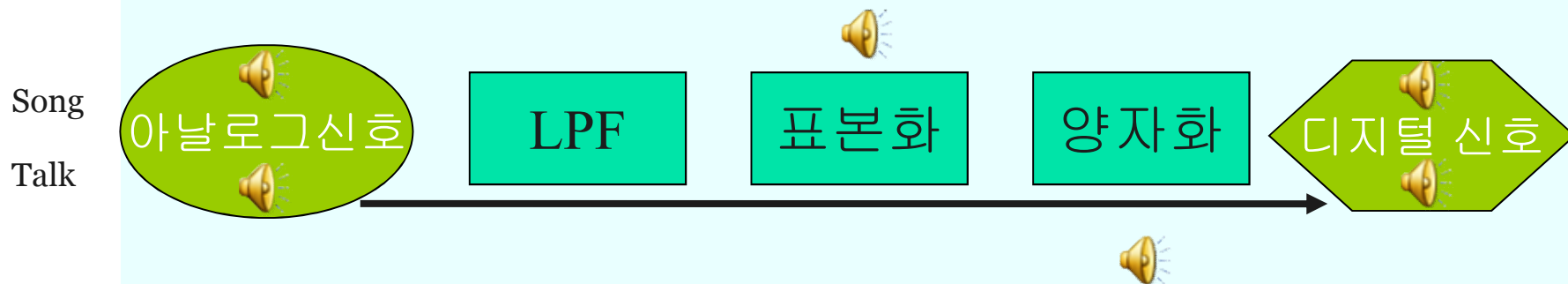
- 아날로그 테이프 녹음 2) 디지털 오디오테이프(DAT) 3) CD-ROM 4) 디지털 디스크
- 주파수 성분의 변화
 - 8KHz로 표본화한 음성신호는 최대 4KHz 까지의 성분만 가짐
- 표본화 주파수
 - 전화음성의 경우 : 8KHz
 - PC : 8KHz, 11.025KHz, 22050KHz, 44.1KHz, 48KHz
 - CSL : 최대 50KHz
 - DAT(Digital Audio Tape) : 44.1KHz, 48KHz
 - CD(Compact Disk) : 44.1KHz
- 해상도(resolution: bit 수)
 - 음성의 경우 12bit 이상이 적당, 보통16bit(65,536 수준의 진폭)로 설정

5.3 양자화(quantization)



■ 신호의 양자화(quantization) 과정

- 신호의 양자화(quantization) 과정; 진폭 축의 디지털화
- 진폭에 관련된 정보를 얼마나 자세히 담을 것인가를 결정하는 것
- 표본화를 통해 쪼개진 값은 연속적인 값을 갖는데, 이 값을 진폭에 따라 연속적이지 않은 각각의 대표값으로 변환하는 과정

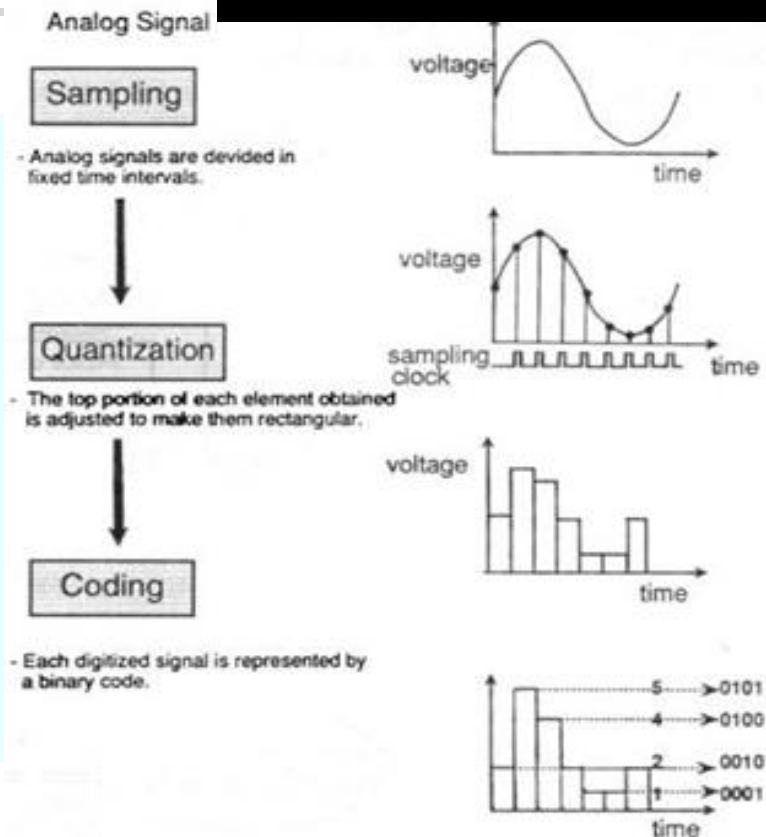


- LPF: 저역통과 필터
- 표본화: Sampling
- 양자화: Quantization

5.3.1 아날로그-디지털 변환 II

(sampling + quantization + coding)

- 아날로그 신호를 디지털 신호로 변환
- A/D 변환 (Analog/Digital)
- 신호의 성분분석에 도움



아날로그신호

LPF

표본화

양자화

디지털 신호

5.4 디지털 신호의 스펙트럼, 스펙트로그램분석

■ 스펙트럼(spectrum)

- 진폭, 주파수 도면으로써 복합파를 분석하여 그 복합파가 어떠한 단순파들의 합으로 이루어진 것인가를 보여주는 방법

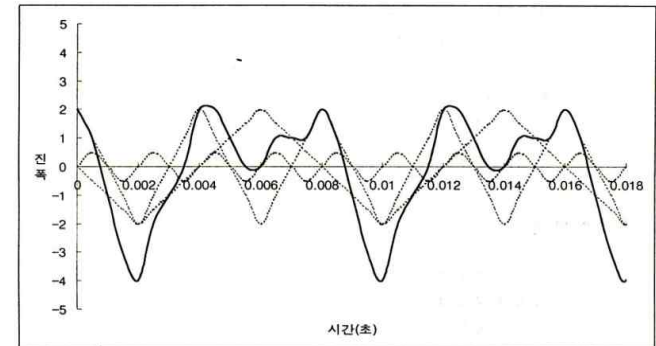
■ 푸리에 변환(Fourier Transform)

- 음성 등의 파형을 기본 주파수(기본음)와 그 정배수의 각 주파수(각 배음)로 분해하는 것; 즉, 어떤 파(波) 중에서 어느 주파수 성분이 얼마만큼 포함되어 있는지를 계산하는 방법
- 모든 파형은 이 푸리에 변환을 통해 기본파와 고주파의 합
- 이산 푸리에 변환(Discrete Fourier Transform; DFT)
- 고속 푸리에 변환(Fast Fourier Transform; FFT)

- 불연속적인 신호를 빠르게 분석하여 스펙트럼으로 나타내게 될 때 사용되는 방식

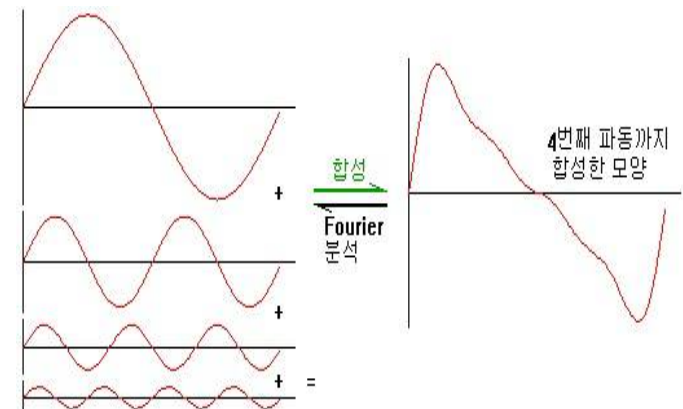
■ 역푸리에 변환(푸리에 합성)

- 기본파와 고주파의 주파수 및 진폭을 알 수 있다면, 이것을 합성해서 하나의 파형으로 나타내는 것



위 그래프에서 점선은 단순파를, 실선은 복합파를 각각 나타낸다.

[그림 7-5] 그림 7-2의 (가), (나), (리)에 보인 세 단순파와 각각 90°, 180°, 0°의 위상차를 가진 세 단순파의 복합으로 만들어진 복합파의 예



5.4.1 디지털 신호의 스펙트럼의 응용



[단독] 보성 연쇄살인사건, 1.2초의 목소리로 범인 잡았다



[쿠키 사회] 119 신고전화에 불과 1.2초 동안 기록된 음성 흔적이 대학 연구팀의 치밀한 분석에 의해 범인의 목소리로 입증됐다.

숭실대 소리공학연구소 배명진 교수 연구팀은 지난 8월 전남 보성에서 남녀 관광객 4명을 연쇄살해한 혐의를 받고 있는 어부 오모(70)씨의 목소리와, 피해자의 119 통화기록에 남은 범인의 목소리를 비교·분석해 오씨의 범행을 입증하는 결정적인 증거를 찾아냈다고 23일 밝혔다. 보성 연쇄살인 사건은 엽기적 범행으로 사회에 상당한 충격을 줬지만 검찰은 지금까지 오씨의 진술 이외에는 목격자나 마땅한 증거자료를 찾지 못해 수사에 어려움을 겪어왔다.

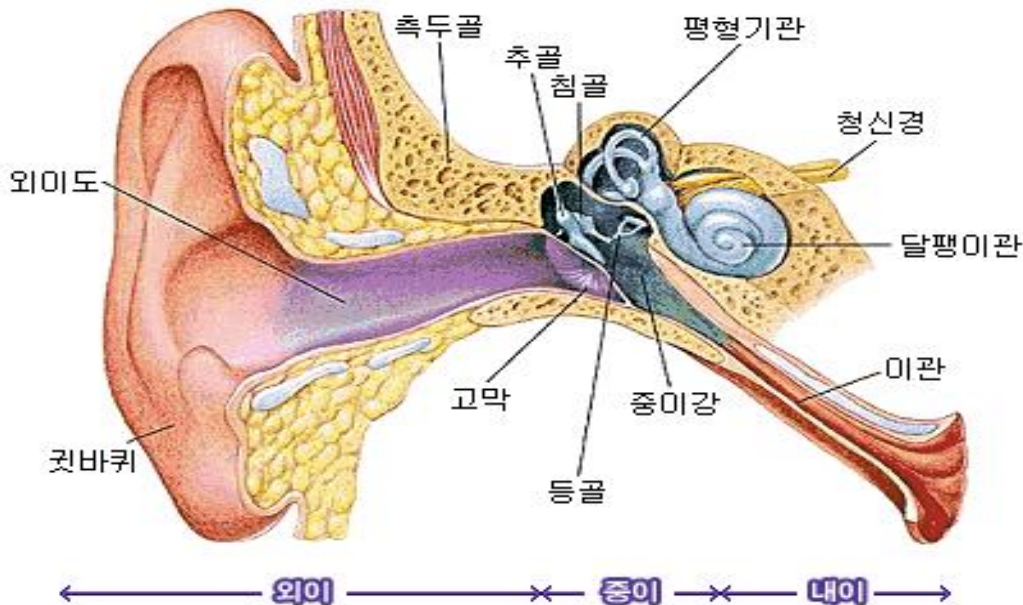
그러나 배 교수팀은 나이가 들면 청력과 발성이 점차 약해져서 소리의 음 대역폭이 줄어드는 성문(聲紋)특성(목소리의 분위기 등 개인 특유의 정보)이 나타난다는 사실에 착안해 목소리 스펙트럼을 비교했다. 그 결과 119 통화에 남겨진 음성은 분석 스펙트럼의 300Hz 근방에서 소리성문(크기, 높낮이, 리듬 등)이 집중적으로 나타났으며, 스펙트럼의 주파수가 점차 감소하는 성문특성을 보였다. 그리고 이 특성은 검찰이 사건 발생 뒤 오씨에게서 녹음한 “어따 전화하고 있어”라는 목소리의 분석 내용과 정확하게 일치했다.

전화기록에 남은 음성이 오씨의 목소리라는 사실이 확인된 것이다. 이와함께 배 교수팀은 119 통화에 녹음된 선박의 엔진소리가 경찰이 나중에 녹음한 오씨 선박의 소리와 일치한다는 점도 밝혀냈다. 배 교수는 “비록 피해자는 유명을 알리했지만 그가 남긴 1.2초의 목소리가 범인을 규명할 강력한 증거로 작용했다”고 말했다. 국민일보 쿠키뉴스 노용택 기자

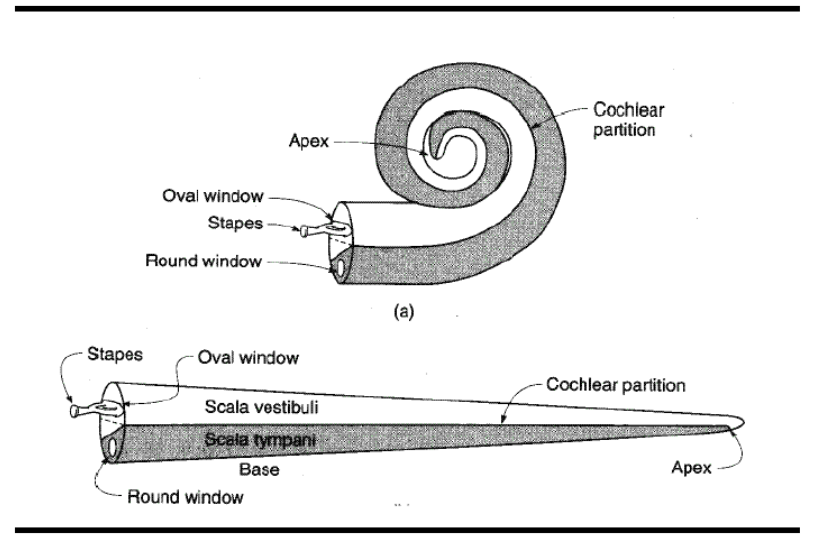
이 사건으로 숨진 여대생 C(20)씨는 납치 과정에서 4차례 119에 전화를 했으며, 이중 4차 통화에 “어디서 무전하니”라는 범인 음성이 고작 1.2초 동안 녹음돼 있었다. 목소리 분석으로 본인 여부를 가리기 위해서는 최소 10분의 분량이 필요해 국립과학수사연구소 조차 119 통화 기록 속의 목소리가 오씨라는 사실을 밝혀내지 못했다.

5.5 인간의 가청영역

- 가청주파수 : **20 – 20,000 Hz**
- 말소리(speech sound)의 대부분은 **5,000Hz 이하**의 정보가 중요
- 모음, 비음, 유음 등은 **5,000Hz 이하**의 정보가 중요하고
- 자음 특히 마찰음과 파찰음은 **5,000Hz 이상**의 주파수 정보가 중요



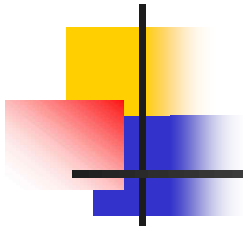
A Partially Uncoiled Cochlea and a Fully Uncoiled Cochlea



5.5.1 Ling's 6 sound test

	성도 공명주파수(Formants)	
	제1공명주파수	제2공명주파수
/m/	250 Hz	
/u/	300 Hz	900 Hz
/a/	700 Hz	1,200 Hz
/i/	250 Hz	2,300 Hz
/sh/	2,000 Hz – 4,000 Hz	
/s/	4,000 Hz – 6,000 Hz	





요약

1. 소리가 일으키는 파동 : 음파(sound wave, 파형(waveform))
2. 스펙트럼
3. 스펙트로그램
4. 음압, 강도, 데시벨(dB)
5. 디지털 녹음과 분석
 1. 디지털과 아날로그
 2. 표본추출률
 3. 양자화
 4. 디지털 신호의 스펙트럼, 스펙트로그램 분석

