1

Sound Source Localization

Joseph B. Keller

Department of Mathematics

Stanford University

Problem: How does an animal determine the location of a sound source from the sound it hears?

This is an *inverse problem.*

Two problems are called inverses of one another if the formulation of each one involves all or part of the solution of the other. One is called the direct problem, the other the inverse problem.

Verbal examples: What is the question to which the answer is

1. Crick?
2. Nine W?
3. Chicken Sukiyaki?
4. Dr. Livingstone I Presume?

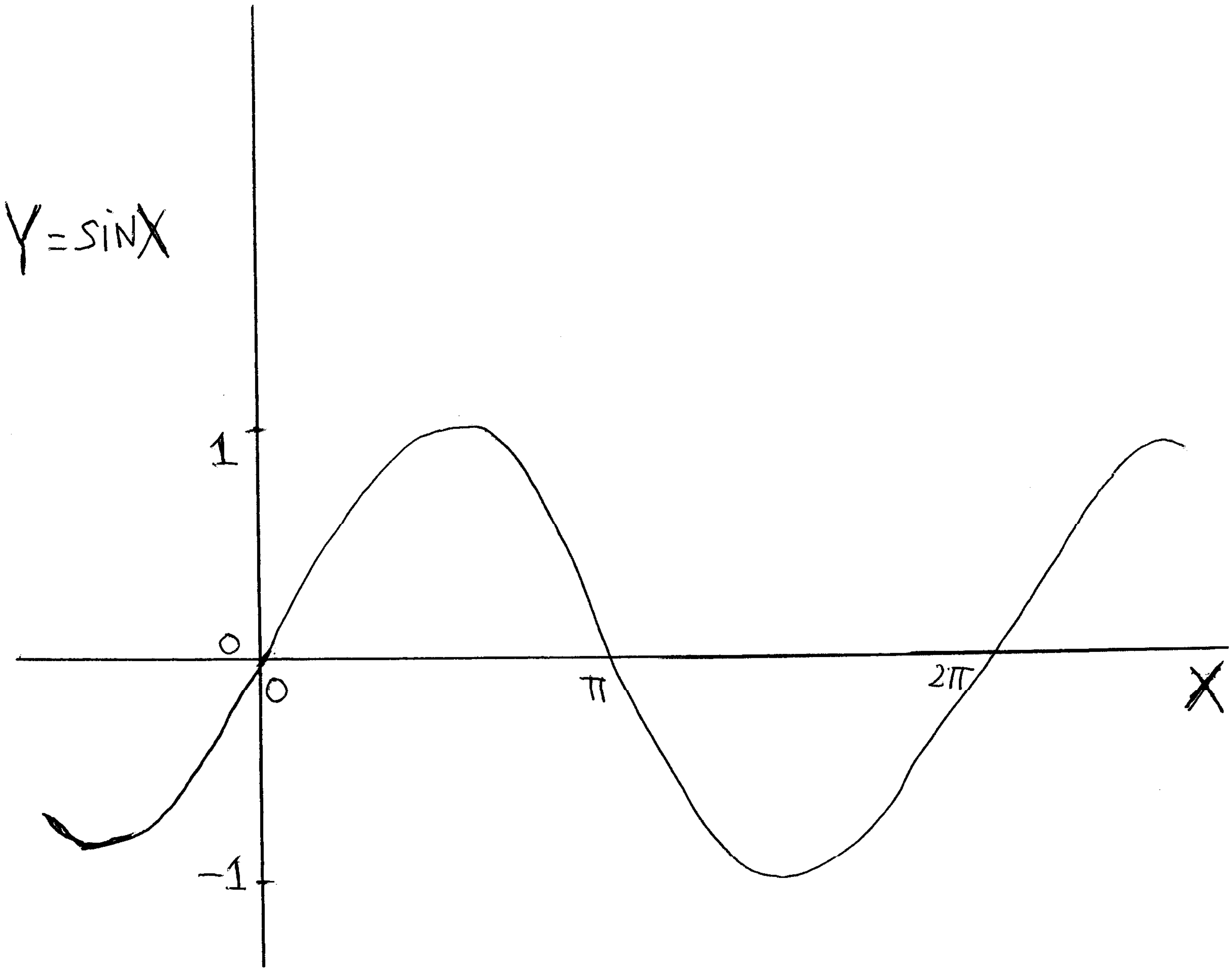
Mathematical Example

Direct Problem: Given x calculate $y = f(x)$.

Inverse Problem: Given y, find $x = f^{-1}(y)$.

Methods of Solution

1. Graphical
2. Tabulation and look-up


Y=sinX
1
0
0
π
2π
X
-1

Direct problem of acoustics: Given source at x_s with time dependence $P(t)$.

Find $P(x, t, x_s)$.

Solution $P(x, t, x_s) = \int_{-\infty}^{t} G(t - t', x, x_s)P(t')\, dt'$

G is impulse response function or time dependent Green's function with animal present

Pressure P_L in left ear at x_L and P_R at right ear at X_R are given by

$$P_L(t, x_s) = \int_{-\infty}^{t} G(t - t', x_L, x_s)\, P(t')\, dt'$$

$$P_R(t, x_s) = \int_{-\infty}^{t} G(t - t', x_R, x_s)\, P(t')\, dt'$$

Inverse problem: Given $P_L(t, x_s)$ and $P_R(t, x_s)$, find x_S and $P(t)$.

To separate the two parts of this problem, we Fourier transform P_L and P_R:

$$p_L\left(\omega, x_s\right) = g\left(\omega, x_L, x_s\right) p(\omega), \qquad p_R(\omega, x_s) = g(\omega, x_R, x_s)\, p(\omega)$$

Divide to eliminate $p(\omega)$:

$$p_R(\omega, x_S)/p_L(\omega, x_s) = g(\omega, x_R, x_s)/g(\omega, x_L, x_s)$$

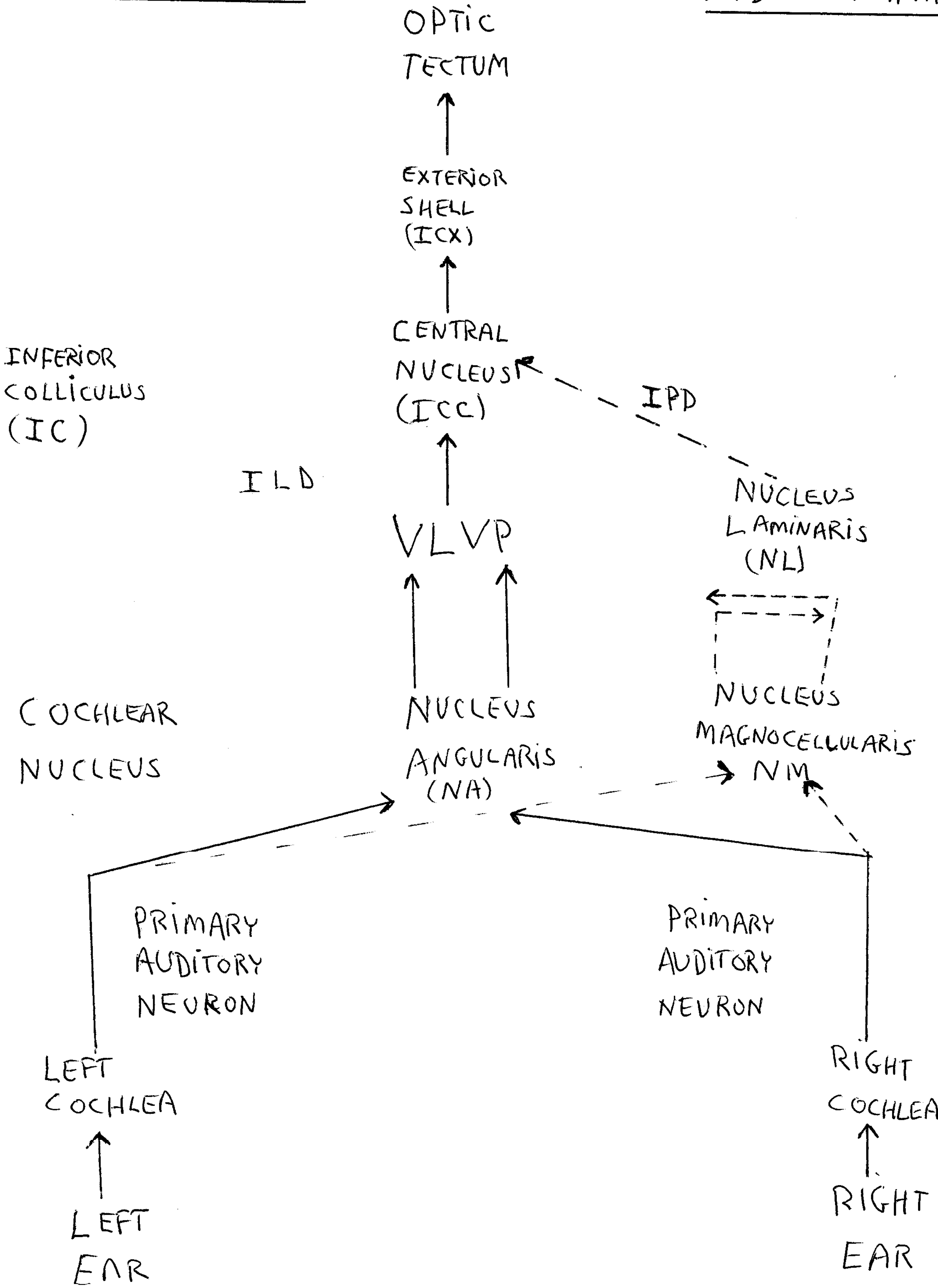
Take logarithms to get two real equations

$$\begin{aligned} ILD(\omega) &\equiv \log\left|p_R\left(\omega, x_s\right)\right| - \log\left|p_L\left(\omega, x_s\right)\right| \\ &= \log\left|g\left(\omega, x_R, x_s\right)\right| - \log\left|g\left(\omega, x_L, x_s\right)\right|. \end{aligned}$$

$$\begin{aligned} IPD(\omega) &\equiv \arg p_R(\omega, x_s) - \arg p_L\left(\omega, x_s\right) \\ &= \arg g\left(\omega, x_R, x_s\right) - \arg g\left(\omega, x_R, x_s\right). \end{aligned}$$

ILD DETERMINATION

IPD DETERMINATION



PRIMARY AUDITORY NEURON FIRES A BURST OF SPIKES IN EACH PERIOD OF ACOUSTIC INPUT.

ONSET OF BURST OCCURS AT A FIXED PHASE OF WAVE.

NUMBER OF SPIKES INCREASES WITH WAVE AMPLITUDE

NA COUNTS NUMBER OF SPIKES PER PERIOD (SPIKE RATE)

TO DETERMINE $\log|p_R|$ AND $\log|p_L|$

VLVP SUBTRACTS TO GET $\log|p_R| - \log|p_L| = ILD(\omega)$

VIA EXCITATORY AND INHIBITORY SYNAPSES

NM LOCATES AND RECORDS BURST ONSET TIME IN EACH CYCLE

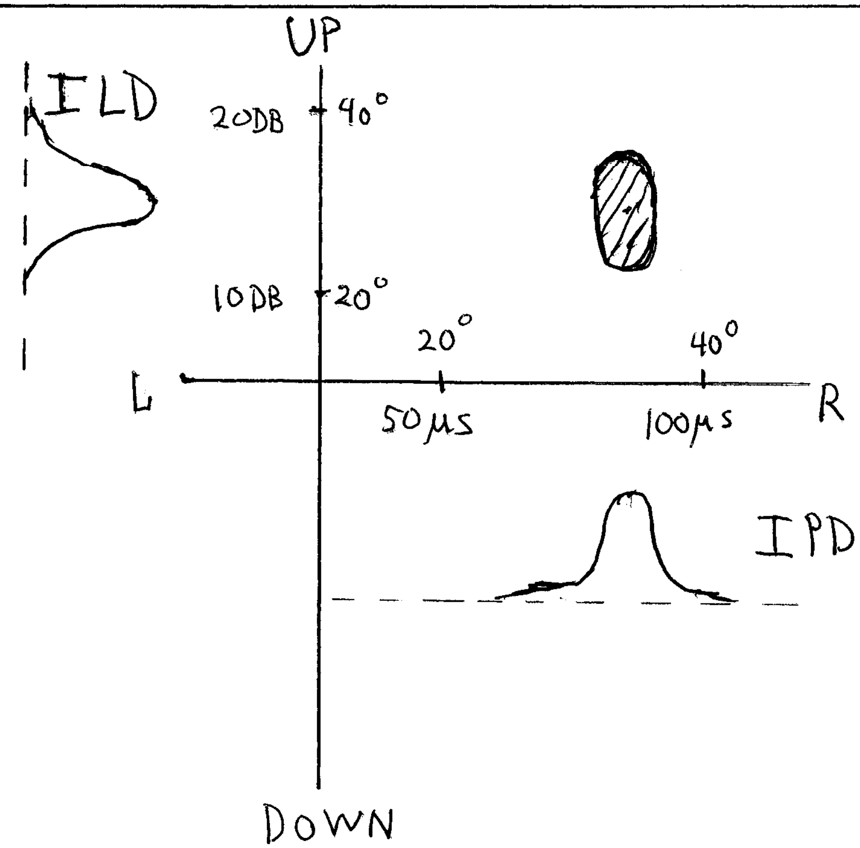
NL ACTS AS A DELAY LINE.

IT CORRELATES RIGHT AND LEFT ONSET TIMES, WITH VARIOUS DELAYS, AND DETERMINES THE DELAY WHICH GIVES THE LARGEST CORRELATION $= IPD(\omega)$.

ICX NEURONS RESPOND TO BOTH ILD AND IPD.

EACH NEURON RESPONDS STRONGEST TO A PARTICLAR $ILD(\omega)_{OPT}$ AND $IPD(\omega)_{OPT}$. THESE OPTIMAL VALUES VARY MONOTONICALLY IN TWO ORTHOGONAL DIRECTIONS ON THE ICX SURFACE.

RESPONSE OF A SPACE-SPECIFIC NEURON IN ICX TO SOUND



BARN OWL USES FREQUENCIES 3KHZ TO 9KHZ.

MOST MAMALS USE FREQUENCIES UP TO 5KH.

AT HIGHER FREQUENCIES, TIMING OF THE ENVELOPE CAN BE USED.

RESPONSE OF A NEURON IN THE OPTIC TECTUM (MODEL)

$$R(\omega) = 1 - \left[\frac{\left(ILD_{OPT}(\omega) - ILD\right)^2}{\alpha(\omega)} + \frac{\left(IPD_{OPT}(\omega) - IPD\right)^2}{\beta(\omega)} \right]^{1/2}$$

FOR EACH NEURON, THE FOUR QUANTITIES

$ILD_{OPT}(\omega)$, $\alpha(\omega)$, $IPD_{OPT}(\omega)$, $\beta(\omega)$

ARE OBTAINED BY FITTING TO THE DATA.

$ILD(\omega)$ MEASURED IN DB, $IPD(\omega)$ IN CYCLES, $-\frac{1}{2} < IPD < \frac{1}{2}$.

$R(\omega) > 0$ EXCITATION

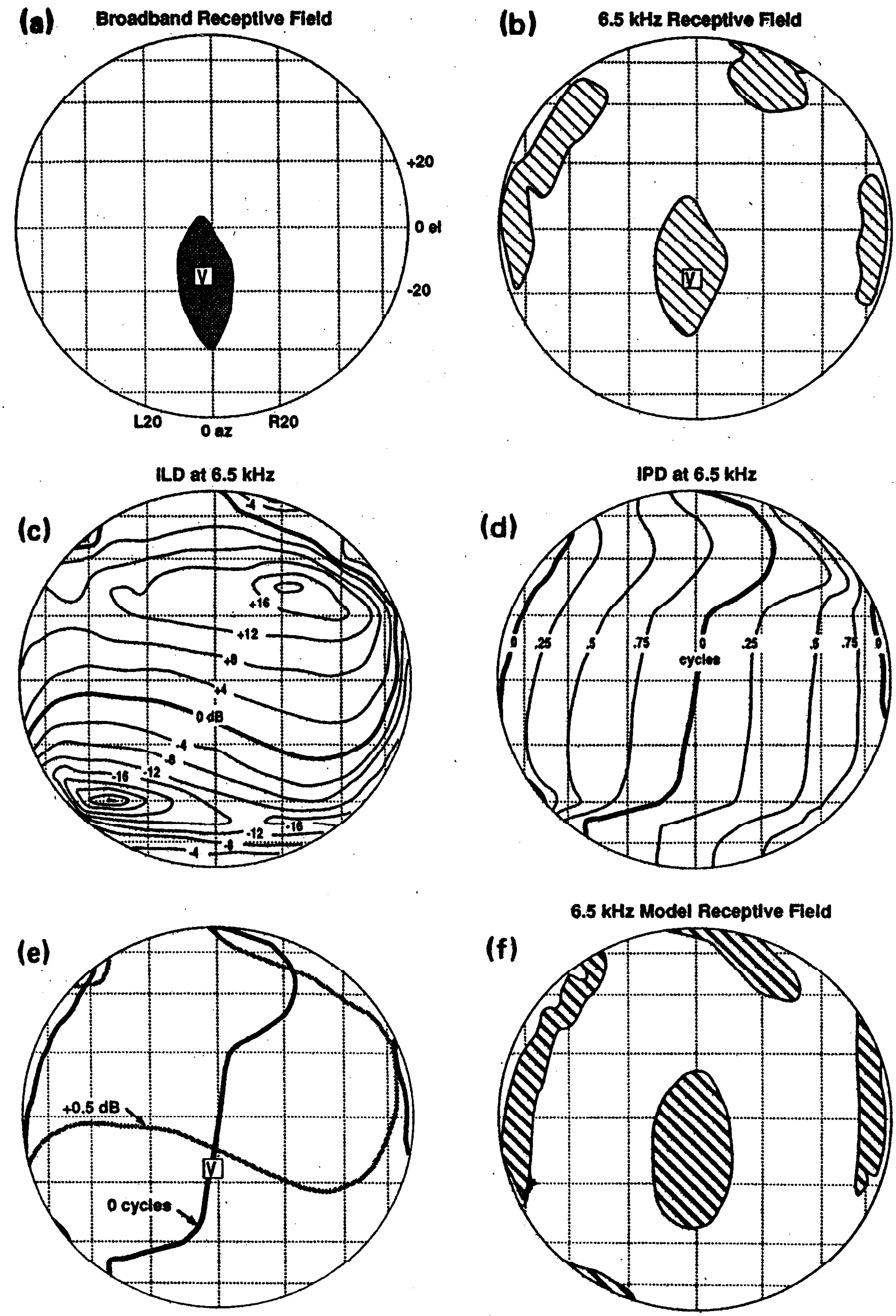
$R(\omega) < 0$ INHIBITION

BROADBAND ACOUSTIC RECEPTIVE FIELD OF NEURON 1

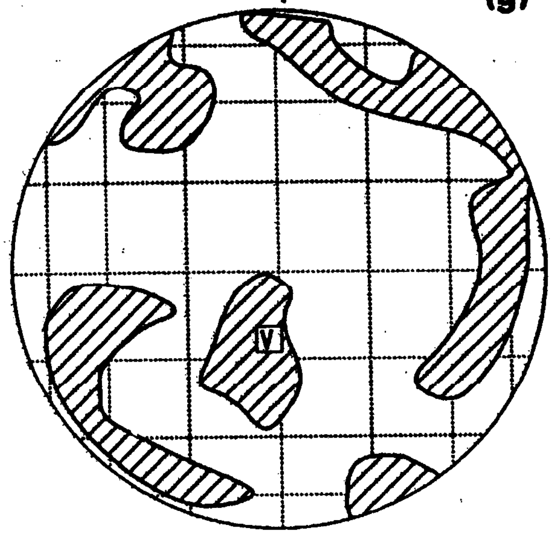
IS CENTERED AT LEFT 2°, DOWN $+19^\circ$

VISUAL RECEPTIVE FIELD OF SAME NEURON 1

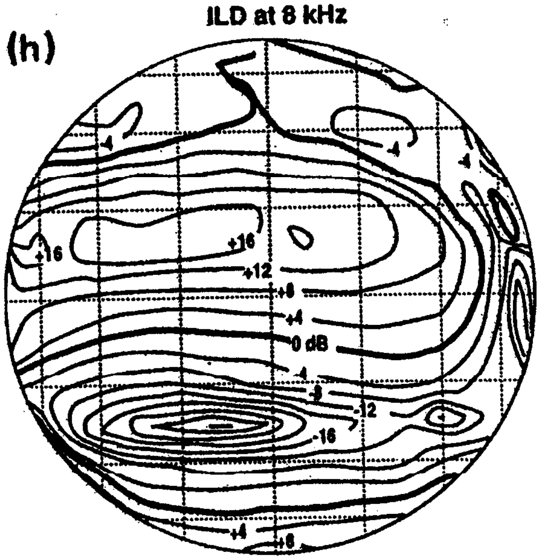
IS CENTERED AT LEFT 2°, DOWN 17°


(a)
Broadband Receptive Field
+20
0 el
-20
V
L20
0 az
R20
(b)
6.5 kHz Receptive Field
V
ILD at 6.5 kHz
(c)
-4
+16
+12
+8
+4
0 dB
-4
-8
-12
-16
-16
-12
-8
-4
IPD at 6.5 kHz
(d)
0
.25
.5
.75
0
.25
.5
.75
0
cycles
(e)
+0.5 dB
V
0 cycles
6.5 kHz Model Receptive Field
(f)

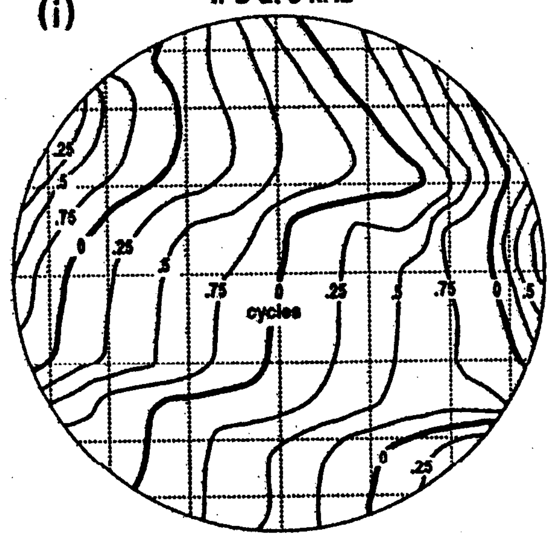
8 kHz Receptive Field (g)



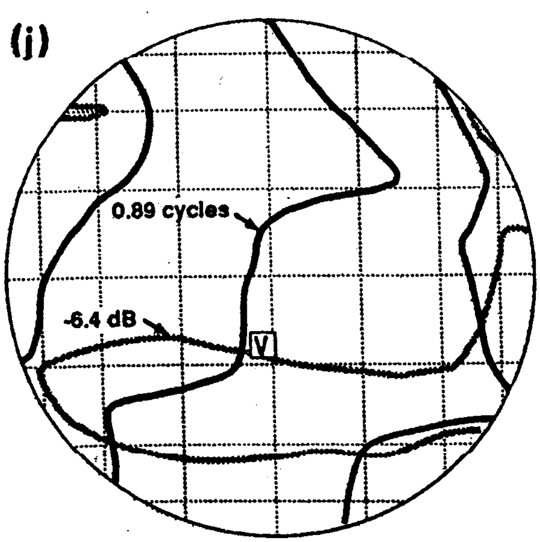
(h) ILD at 8 kHz



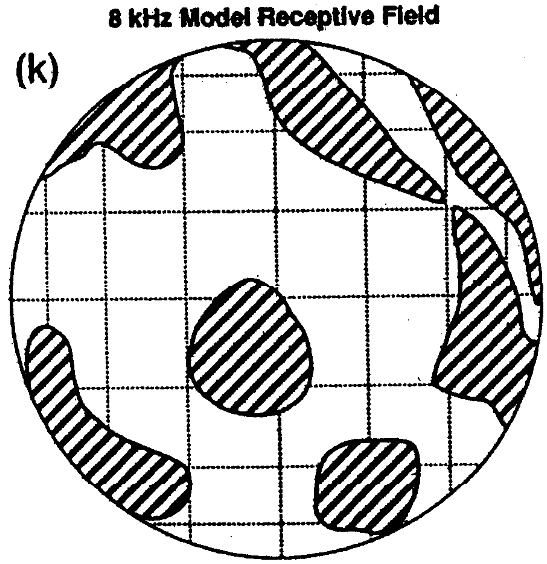
(i) IPD at 8 kHz

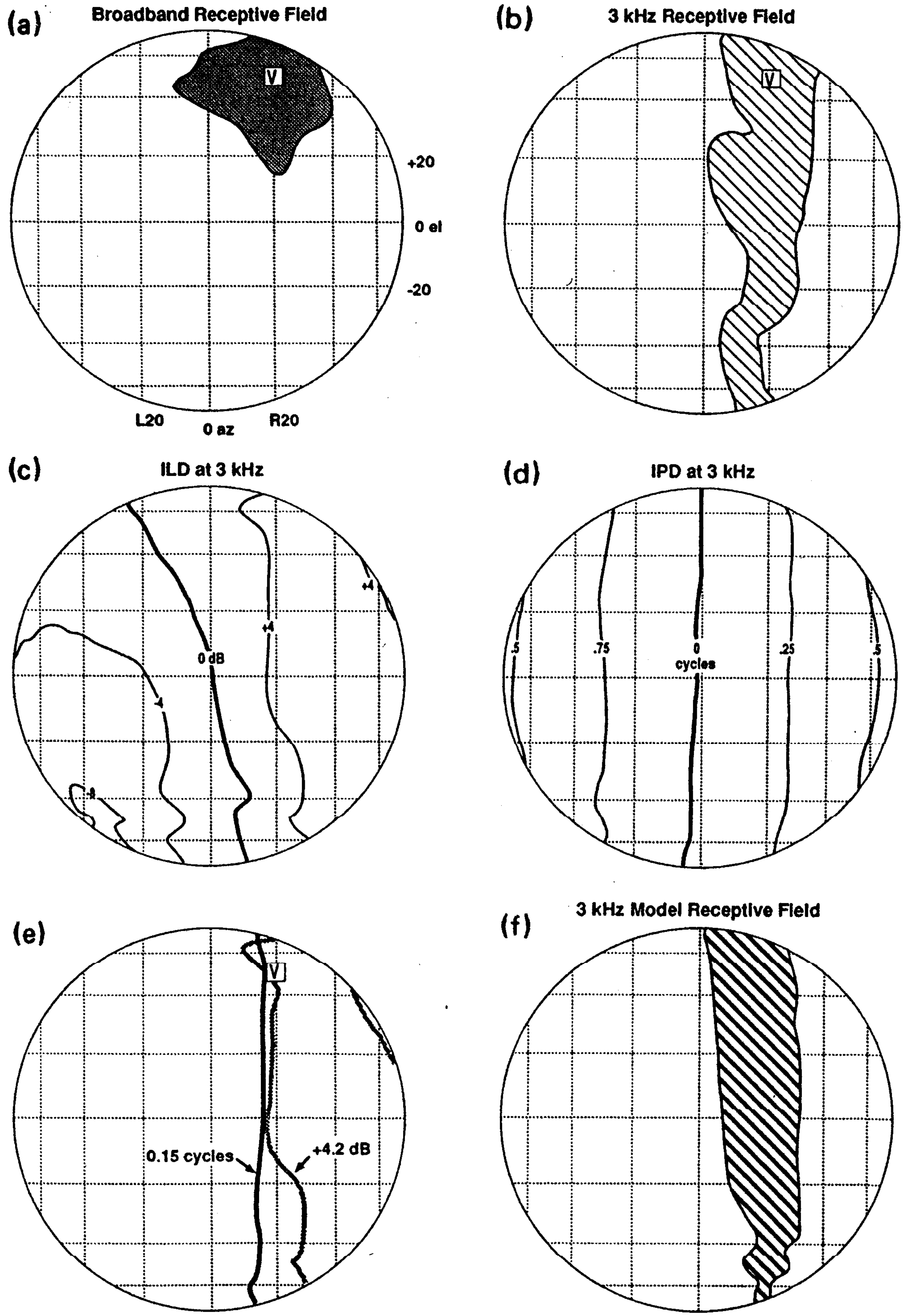


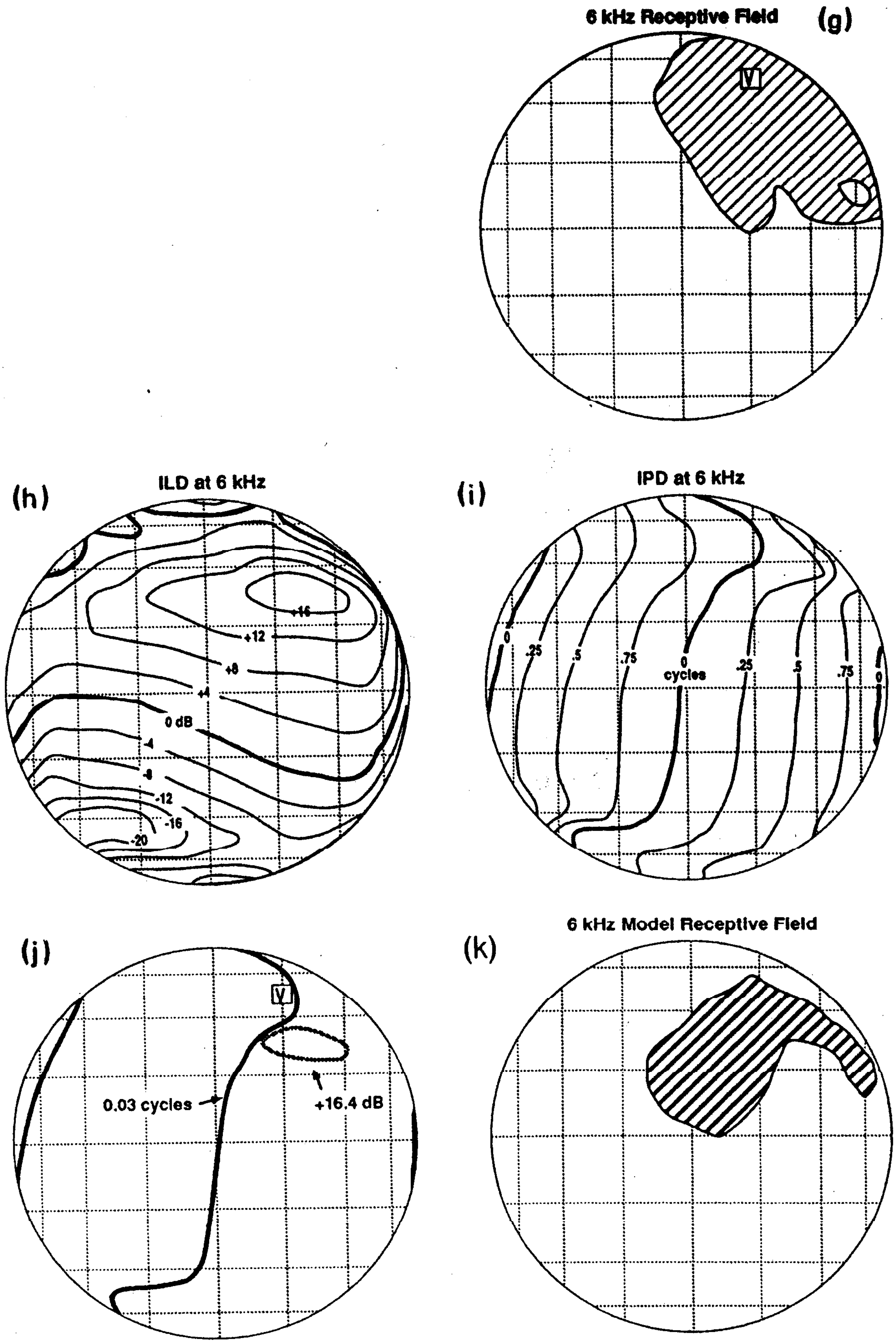
(j)

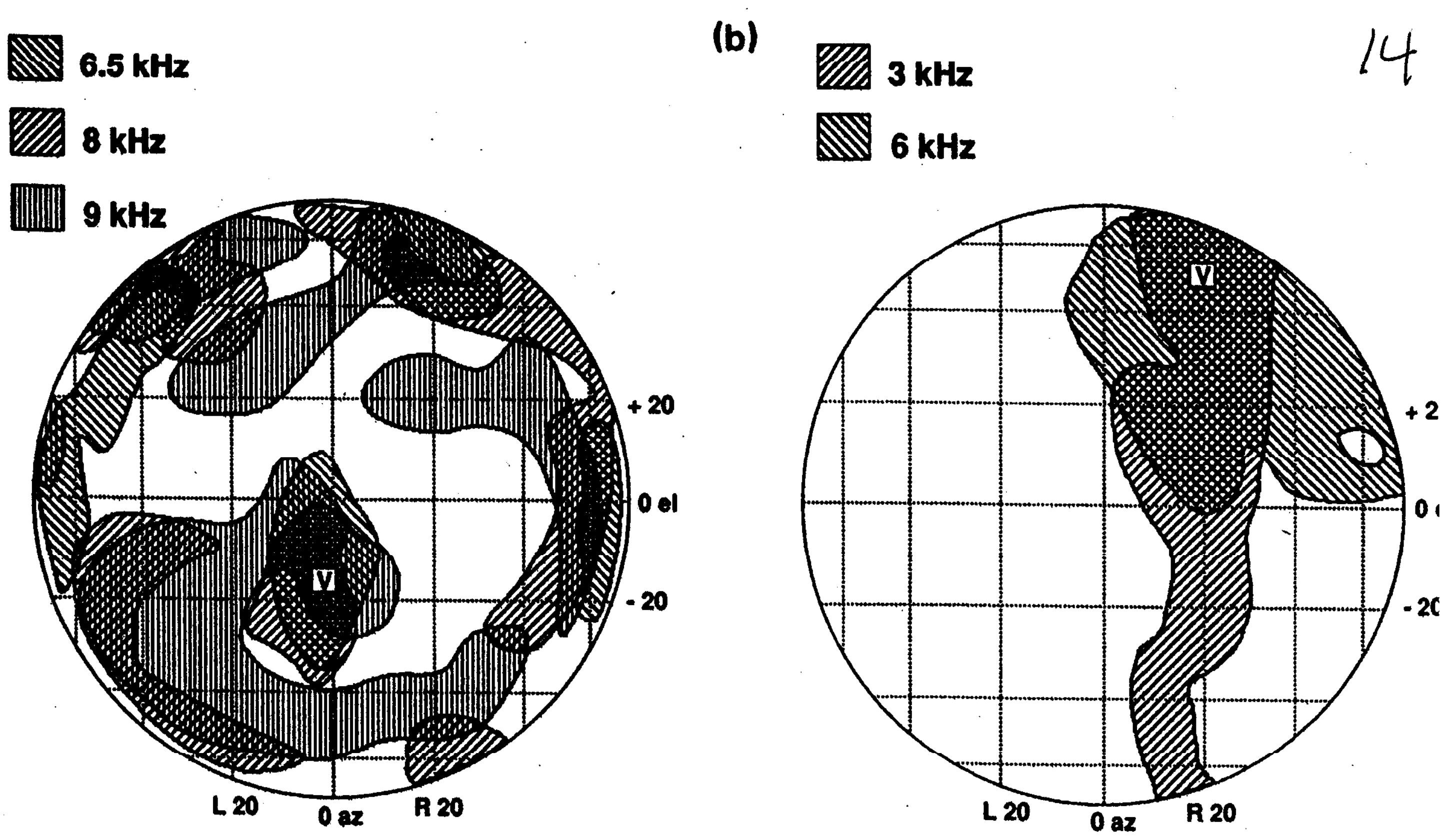


(k) 8 kHz Model Receptive Field

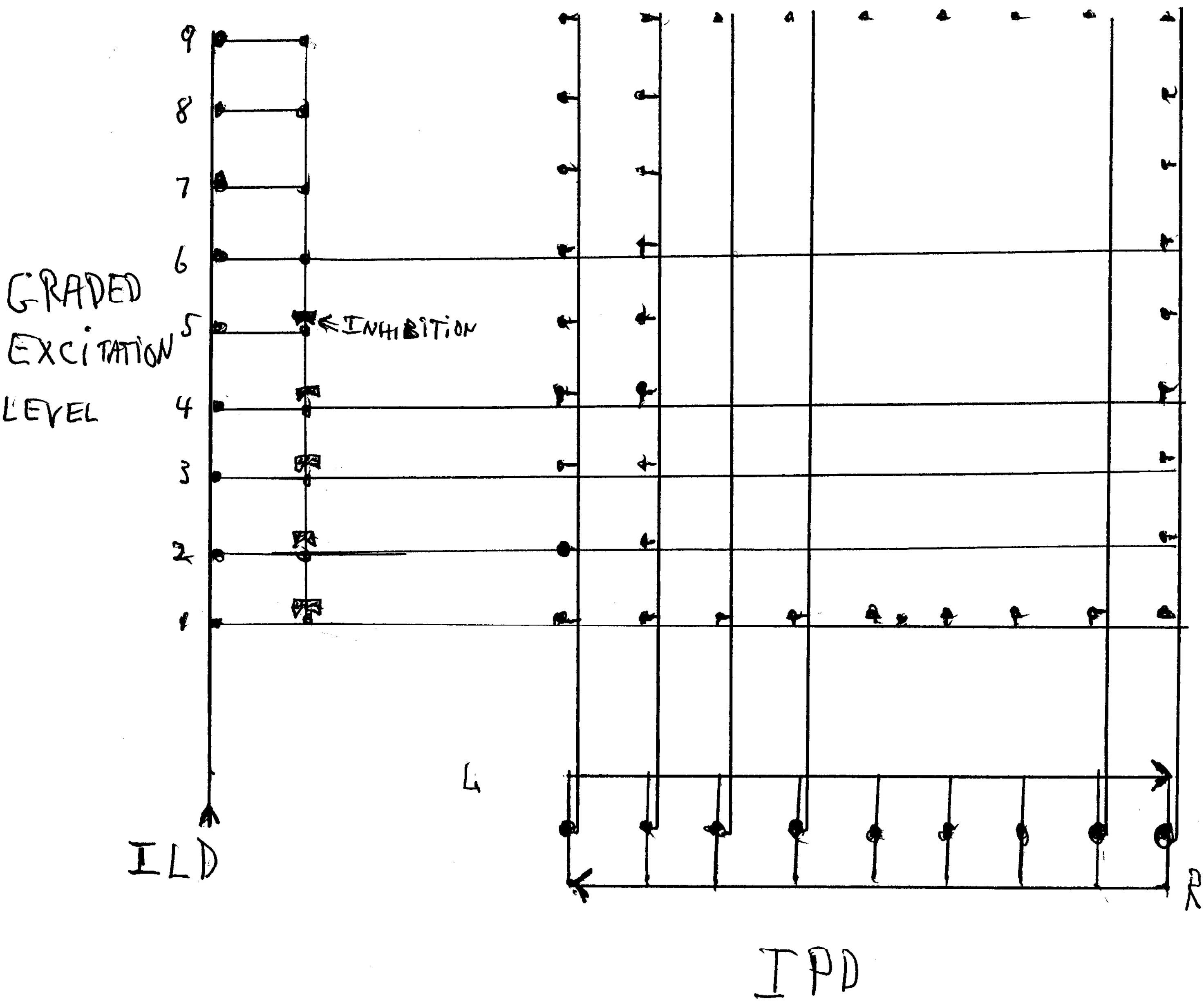



(a)
Broadband Receptive Field
V
+20
0 el
-20
L20
0 az
R20
(b)
3 kHz Receptive Field
V
(c)
ILD at 3 kHz
0 dB
+4
+4
-4
-8
(d)
IPD at 3 kHz
.5
.75
0
cycles
.25
.5
(e)
V
0.15 cycles
+4.2 dB
(f)
3 kHz Model Receptive Field


6 kHz Receptive Field
(g)
V
(h)
ILD at 6 kHz
+16
+12
+8
+4
0 dB
-4
-8
-12
-16
-20
(i)
IPD at 6 kHz
0
.25
.5
.75
0
cycles
.25
.5
.75
0
(j)
V
0.03 cycles
+16.4 dB
(k)
6 kHz Model Receptive Field

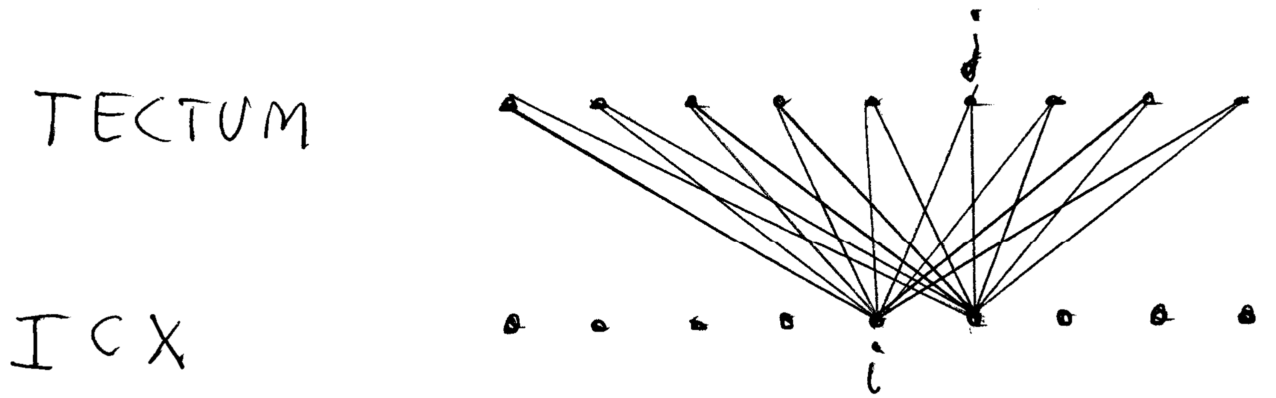

6.5 kHz
8 kHz
9 kHz
V
+ 20
0 el
- 20
L 20
0 az
R 20
(b)
3 kHz
6 kHz
V
L 20
0 az
R 20

ICX EXCITATION



SINGLE FREQUENCY LAYER

OPTICAL TECTUM EXCITATION



EACH ICX NEURON EXCITES TECTUM NEURONS WITH DIFFERENT SYNAPTIC STRENGTHS w_{ij}.

w_{ij} VARY DURING DEVELOPMENT (FIRST 200 DAYS).

CAPABILITY OF VARYING IS CALLED PLASTICITY.

DEVELOPMENT CAN BE ALTERED BY

a) EAR PLUGS

b) PRISM GLASSES.

RESULT IN NEW AXONS AND MODIFIED SYNAPTIC STRENGTHS.

DEVELOPMENT

SUPPOSE ICX NEURON i SIGNALS TECTAL CELL j TO FIRE.

SUPPOSE OPTICAL SYSTEM SIGNALS TECTAL CELL j THAT THERE IS A MOUSE IN VISUAL RECEPTIVE FIELD OF TECTAL CELL j. THEN w_{ij} IS INCREASED.

IF THERE IS NO MOUSE, w_{ij} IS DECREASED.

THERE IS A MAXIMUM VALUE OF w_{ij} (SATURATION)

NEGATIVE VALUES DENOTE INHIBITION.

WHY DO OWLS EAT MICE?

REFERENCES

ERIC I. KNUDSEN ET AL ANN. REVS. NEUROSCI. 18 19-44, 1995

" NATURE 412, 322-328, 2002

M. KONICHI, COMP BIOCHEM PHYSIOL. A 126 459-469, 2000