UNIVERSITY OF CALIFORNIA,
IRVINE

# Five Studies in Measurement and Psychophysics

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Social Science

by

Christopher William Doble

Dissertation Committee:
Professor Bruce Berg, Chair
Professor Jean-Claude Falmagne
Professor Louis Narens

2002

UMI Number: 3067696

UMI®

UMI Microform 3067696



ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

The dissertation of Christopher William Doble
is approved and acceptable in quality
and form for publication on microfilm:

Committee Chair

University of California, Irvine
2002

# DEDICATION

In loving memory of my father.

William L. Doble

# Contents

# List of Figures

# List of Tables

## ACKNOWLEDGEMENTS

I have received generous financial support from the UCI School of Social Sciences (through a Merit Fellowship), the Institute for Mathematical Behavioral Sciences, and the sons and friends of Justine Lambert. I would like to thank Kluwer Academic Publishers for permission to include Chapter 1 of this dissertation, which was originally published in the journal *Order*. I would also like to thank the American Psychological Association for permission to include Chapter 4 of this dissertation, which is to be published in the journal *Psychological Review*.

I would like to express my heartfelt thanks to my advisor, Jean-Claude Falmagne. His contagious enthusiasm for science and mathematics and his unfailing support of my academic endeavors are the chief reasons this dissertation is complete. It was an honor to be his student the last three years, and it is an honor now to call him my friend.

I also would like to thank the faculty and staff in UCI's Department of Cognitive Sciences. I gained a great deal from my associations with such world-class researchers as Bruce Berg, Geoff Iverson, Charlie Chubb, Louis Narens, Bill Batchelder, Duncan Luce, Barbara Dosher, Ece Batchelder, and many others. I have received much support from my office mate Yung-Fong Hsu, from the IMBS Administrator Janet Phelps, from Social Sciences Graduate Counselor Kathy Alberti, and from all the staff at UCI who have been so helpful and encouraging. I am especially indebted to my neighbor and friend Lin Nutile, who selflessly has provided assistance and inspiration every day for the past two years.

It was an honor for me to co-author a paper with Jean-Paul Doignon and Peter Fishburn. I also thank Janos Aczél and David Eppstein and for insight and support.

I am fortunate to have the unconditional love of a remarkable woman, my mother, Anita Fellers. I also am blessed with an extraordinary sister, Lisa, and a wonderful grandma and aunt, all of whom have inspired me countless times. My dear friend Rose Mendoza will never know how important she has been to me, and this thank you is hardly enough.

This work is dedicated to my father, in loving memory of whom I do all things.

## CURRICULUM VITAE

### Christopher William Doble

| | |
|---|---|
| 1992 | A.B. Psychology. Stanford University. Stanford. CA |
| 1999 | M.S. Mathematics. California Polytechnic State University. San Luis Obispo. CA |
| 2002 | Ph.D Social Science. University of California. Irvine. CA |

## ABSTRACT OF THE DISSERTATION


Five Studies in Measurement and Psychophysics

by

Christopher William Doble

Doctor of Philosophy in Social Science

University of California, Irvine, 2002

Professor Bruce Berg, Chair



The studies reported in this work, which are rather diverse in character, are linked by a common use of mathematics in the modeling of behavioral and physical phenomena. The first two studies report mathematical results which were inspired by the modeling of the evolution of psychological "states"—states of preference, states of knowledge, etc.—via stochastic processes on combinatoric structures. The first study details a mathematical investigation of particular types of order relations dubbed 'almost-connected orders,' which are shown to possess properties which naturally suggest their incorporation into stochastic models of preference evolution. The second study, inspired in part by practical problems in the modeling of states of knowledge, briefly examines two approaches for systematically generalizing 'partial orders.' The third study contains an investigation of two measurement-theoretic properties of invariance, termed 'meaningfulness' and 'dimensional invariance,' which have been used in the search for functions which may be said to relate empirical variables in a "lawful" way. The fourth and fifth

studies detail results derived from mathematical and empirical investigations of a phenomenon in psychoacoustics called the 'near-miss to Weber's law.' It is shown that the parameter estimates typically obtained for a customary model of this phenomenon are inconsistent with a common averaging over experimental conditions, giving an obvious warning against the use of the model. An alternative model is shown to provide a good fit to well-known data and to data recently collected in our laboratory. These latter data suggest a systematic covariation of the parameters in the alternative model consistent with a 'gain-control' description of auditory intensity discrimination.

# Introduction and Preview

Though the title of this work, "Five Studies in Measurement and Psychophysics," may suffer from a certain lack of panache, it nonetheless conveys (perhaps implicitly) the accurate notion that the studies reported in the work cover multifarious topics which may not be immediately linked. The studies, described in five chapters, each meant for separate publication, touch upon diverse areas of the behavioral sciences and the philosophy of science. If there is a common link among the studies, it is that approaches to modeling in these areas have been, historically, decidedly mathematical. The approach in this work is no different.

The first two chapters detail mathematical results motivated by work in the modeling of the evolution of psychological "states" via combinatoric structures. For the work in Chapter 1, these states correspond to states of preference among a finite set of alternatives, and the models belong to a class of stochastic models formulated by Falmagne and colleagues (Falmagne, 1997; Falmagne and Doignon, 1997; Falmagne and Ovchinnikov 2002) and successfully applied by Regenwetter et al. (1999) and Hsu et al. (to be submitted). In these models, preferences are presented via orderings of the alternatives, and the evolution of preferences is depicted as a random walk on the orderings. Under very general axioms for these random walks, strong results regarding the asymptotic behavior of the stochastic processes may be derived (e.g., Theorems 1-5 in Regenwetter et al., 1999). Using these results and the parameter estimates obtained from the application of these models, one may be able to draw conclusions regarding the flow of information effecting the relevant changes of preference.

Application of these models may entail the use of a particular type of order relation called a 'weak order' (see Definition 1.3). Essentially a ranking of the alternatives with ties allowed, a weak order is the order type elicited in opinion polls which require respondents to give numerical ratings of each alternative. Such ratings are immediately coded as weak orders, and a model which assumes a random walk on weak orders may naturally be applied to the data. Such a model has been used, for instance, in the analysis of presidential election opinion poll data (Regenwetter et al., 1999; Hsu et al., to be submitted).

Though respondents necessarily give weak order rankings when polled, it could be that intermediate preference states, i.e., the states between the polls, are less constrained than weak orders. With this in mind, one might seek to classify the "bridge" relations between weak orders which could give plausible representations of an individual's preference state. A goal of the mathematical study in Chapter 1 is to classify order relations which could be said to lie "between" two weak orders, that is, to classify the relations $R$ such that $W \subset R \subset W'$, where $W$ and $W'$ are weak orders such that $W \subset W'$ with no weak order $W''$ that satisfies

$W' \subset W''' \subset W''$. It is proved in Theorem 1.15 that such an $R$ must satisfy the axiom

(1) if $xRy$ and $yRz$, then $xRw$ or $wRz$

for all $x$, $y$, $z$, $w$ in the ground set. Such relations are called in the Chapter 'almost connected orders,' or 'ac-orders,' because (1) is a natural generalization of a connectedness condition. Chapter 1 contains an examination of ac-orders and their relationship to other order relations, especially weak orders. It turns out that every ac-order is bracketed in a natural way by two weak orders (see Theorem 1.17 and Lemma 1.21). One of the weak orders, called the 'contraction' weak order, is the maximum in the set of weak orders included in the ac-order. The other, called the 'height' weak order, is minimal, but not necessarily the minimum, in the set of weak orders that include the ac-order.

Central to the formulation of the aforementioned models of preference evolution are conditions involving movement from one member of a family of order relations to another member of the family. One strong condition, called 'wellgradedness' (Doignon and Falmagne, 1997), allows movement in an efficient manner: a family $\mathcal{F}$ of relations on a set $\mathcal{Y}$ is *well graded* if, for any relations $A, B \in \mathcal{F}$ there exists a finite sequence of relations $A = F_0, F_1, \ldots, F_{|A \triangle B|} = B$ in $\mathcal{F}$ such that $|F_{i-1} \triangle F_i| = 1$, for $i = 1, \ldots, |A \triangle B|$ (where $\triangle$ stands for the symmetric difference between sets). It is shown in Theorem 1.29 that the family of all ac-orders on a finite set $\mathcal{Y}$ is well graded if, and only if, $|\mathcal{Y}| \leq 4$. However, for any finite $\mathcal{Y}$, the family of all ac-orders on $\mathcal{Y}$ is necessarily 'downgradable': any nonempty ac-order on $\mathcal{Y}$ can be trimmed down by removing pairs one by one. until the empty order is reached. without ever leaving the family of all ac-orders on $\mathcal{Y}$ (Theorem 1.32). Also. the family is 'upgradable': any ac-order which is not a linear order may be enlarged by adding pairs one by one. until a linear order is formed. without ever leaving the family (Theorem 1.32). Such results are important in the application of this family of relations to the random walk models of preference evolution.

The second chapter also describes work motivated by the modeling of psychological "states." this time with the states corresponding to states of knowledge. Consider. for example. a situation in which a school official is to assess a student's mastery of an academic subject. such as high school algebra. One way to describe the student's mastery is to associate with the student a subset of the set of all types of high school algebra problems. with the subset—called the student's "knowledge state"—corresponding to the problems that the student is able to solve. As the student learns. the subset necessarily changes. expanding as the student's knowledge of the subject expands. The mathematical underpinnings of such a situation are well studied in Doignon and Falmagne (1999). Their work (along with the work of several others: see Doignon and Falmagne. 1999 for references) has resulted in the development of successful. on-line. assessment and instructional systems for several academic subjects.

A key component of these assessment and instructional systems is an underlying structure of problems and their prerequisite problems. This structure may be represented by a family $\mathcal{K}$ of subsets of a finite set $Q$. The elements of $Q$, which correspond to the problems to be mastered, are called *items*, and members of $\mathcal{K}$ are called *states*. An important practical issue is the construction of such a structure, that is, the construction of a plausible family $\mathcal{K}$ which gives states that a student may occupy on the way to mastery of the academic subject. Since this family may be too large to be listed explicitly, it often is arrived at indirectly, via experts' responses to questions of the form

> If a student has failed to solve all of the problems in the set $A$, will she also fail to solve problem $q$?

for all nonempty subsets $A$ of $Q$, and all items $q$ in $Q$. The responses define a relation $P$ from $2^Q \setminus \{\emptyset\}$ to $Q$, with the pair $(A, q)$ being in the relation precisely when the response to the above question is "Yes." A family $\mathcal{K}'$ of subsets of $Q$ may be derived from $\mathcal{P}$ by the equivalence

$$(2) \qquad K \in \mathcal{K}' \iff (\forall (A, q) \in \mathcal{P} : A \cap K = \emptyset \Rightarrow q \notin K).$$

If only singletons $A$ are used, then the family $\mathcal{K}'$ corresponds (after a recoding) to a partial order on the items in $Q$. (This result is due to Birkhoff, 1937.) Such a family has the property that each item is contained in a unique minimal set, or 'background,' of $\mathcal{K}'$. Thus, if the original family $\mathcal{K}$ contains an item with more than one background, then a relation $\mathcal{P}$ containing only singletons $A$ will not 'recover' $\mathcal{K}$, in that $\mathcal{K}'$ will not equal $\mathcal{K}$ (see Definitions 2.2).

The work in Chapter 2 is motivated by the following questions: If the original family $\mathcal{K}$ has items with more than one background. what can be said about relations $\mathcal{P}$ which recover $\mathcal{K}$? In particular. what can be said about

$$k = \min_{\substack{\mathcal{P} \\ \mathcal{P} \text{ recovers } \mathcal{K}}} \max\{|A| : (A, q) \in \mathcal{P}\}\ ?$$

Does $k$ always equal the maximum number of backgrounds in $\mathcal{K}$. as it does when this number is one? Answers to such questions develop a link between the querying of experts and the types of structures which may be recovered from this querying. This link has obvious practical implications. In addition. the link may allow a systematic. mathematical generalization of partial orders. with the generalization following a progression from families whose items have at most one background (i.e.. partial orders) to families whose items have $2, 3, 4, \ldots, n$ backgrounds. Chapter 2 contains results which specify this link.

Chapter 3 comprises a measurement-theoretic investigation of two properties of invariance of possible scientific laws. It is standard to require that scientific laws be invariant in form under certain transformations of the relevant objects. especially transformations involving equivalent representations of the objects using different measurement scales (see e.g. Narens. 2002). Of course. terms such as "invariance." "form." and "scientific law" are inexact and should be defined carefully. The definitions used in Chapter 3 follow closely those of Falmagne and Narens (1983). and these authors' terms are used for the two types of invariance considered in the chapter. namely. 'meaningfulness' and 'dimensional invariance' (see Definitions 3.5 and 3.6). The main result of the chapter. which gives insight into the relationship between the two formulations of invariance. generalizes a result by these authors. The generalization lies in the types of transformations for which invariances are considered. It is common in the measurement literature for invariances to be considered for changes in representation of the variables giving the results of measurement, i.e., for certain strictly increasing, surjective transformations that act individually on the variables. However, it turns out that there are important cases in which invariance holds under transformations that are not 'factorizable,' that is, under transformations that cannot be written as functions on separate variables. The Lorentz transformation in physics is an example. It turns out that, in the general setting considered in Chapter 3, extensions of Falmagne and Narens' invariance formulations stand in the same relationship as in the original setting, namely, the two formulations are equivalent under a natural condition relating members of the family of functions under consideration (see Definition 3.10 for this condition).

The two formulations are independent, however, and there exist physical laws which satisfy meaningfulness but not dimensional invariance (see Example 3.1). Examinations of physical laws which are not dimensionally invariant, of whether these laws allow associated formulations which are dimensionally invariant, and of how those associated formulations are obtained are examined briefly. These examinations suggest the use of dimensional invariance beyond the typical use in classical physics, i.e., beyond the method of dimensional analysis.

This study of invariance is motivated by the characterization of functions which may be said to relate empirical variables in a "lawful" way. Such a characterization is sought via examination of putative invariances of the measurement theories of the variables, which may greatly constrain the possible forms of empirical laws (e.g., Luce, 1959, 1964, 1990; Luce et al., 1990; Osborne, 1970; Falmagne and Narens, 1983; Aczél et al., 1986; Kim, 1990). Even if a relation is considered lawful, though, care must be taken in applying the relation as a model of an empirical situation involving further invariance. For instance, in modeling certain psychophysical data, an important but perhaps overlooked invariance is robustness of the model to averaging—averaging over subjects, over experimental trials, over experimental conditions, etc. (See, for example, Heathcote et al., 2000.) Considerations of robustness to averaging may help eliminate candidate models which otherwise seem appropriate, as illustrated by the following example, the detailed discussion of which comprises much of Chapter 4.

The power law, ubiquitous in psychophysical modeling, has been used to describe many data which deviate from Weber's law (cf. Baird and Noma, 1978). Weber's law holds when the ratio $\frac{\Delta(x)}{x}$ is constant, where $\Delta(x)$ is the smallest perceptible positive difference between two stimuli with intensities $x$ and $x+\Delta(x)$ (Fechner, 1860). In many empirical situations, including judgments of line lengths (Guilford, 1932; Hovland, 1938), discriminations of light intensities (Mansfield, 1976), and discriminations of pure-tone intensities (e.g. Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988), the fraction $\frac{\Delta(x)}{x}$ decreases with increasing $x$

in such a way that the model

(3) $$\Delta(x) = Cx^{\alpha}$$

provides a good fit to the data, with estimates of $\alpha$ typically less than 1 and greater than about .5. This has been termed the 'near-miss to Weber's law,' since Weber's law holds when $\alpha = 1$ (McGill and Goldberg, 1968a,b).

Since the definition of $\Delta(x)$ as the "smallest perceptible positive difference" is ambiguous, it is helpful to make explicit reference to the criterion for discrimination. To this end, let $\xi_{\nu}(x)$ be the stimulus intensity judged greater than intensity $x$ with probability exactly equal to $\nu$, and let $\Delta_{\nu}(x) = \xi_{\nu}(x) - x$ (cf. Luce and Galanter, 1963; Falmagne, 1985). With this notation, Weber's law is expressed by the equation

(4) $$\Delta_{\nu}(x) = C(\nu)\, x,$$

in which the constant of proportionality $C(\nu)$ is strictly increasing with $\nu$. Values adopted for the discrimination criterion $\nu$ typically fall between .70 and .80, with no universal convention (cf. McGill and Goldberg, 1968a,b; Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988; Schroder et al., 1994; Neff and Jesteadt, 1996). Equation (3) may then be written as

(5) $$\Delta_{\nu}(x) = C(\nu)\, x^{\alpha(\nu)}$$

to indicate the possible dependence of $C$ and $\alpha$ on $\nu$.

As a standard practice in psychoacoustic (and other psychophysical) intensity discrimination experiments, data are averaged over order of stimulus presentation in a two-alternative, forced-choice task (or over position in a visual discrimination task). This averaging enforces a particular condition, called by Falmagne (1985) the *balance condition*, on the function $\xi_\nu$. The balance condition is equivalent to the equation

$$\xi_\nu[\xi_{1-\nu}(x)] = x \tag{6}$$

holding for all intensities $x$ and all criteria $\nu$ such that $0 < \nu < 1$. It is proved in Chapter 4 that the invariance given by (6) imposes a powerful mathematical constraint on the values for the exponent $\alpha$ in the near-miss equation given by (5): under the balance condition, the exponent $\alpha(\nu)$ in (5) necessarily equals 1 for all $\nu$ (see Theorem 4.1). For the many data which give an estimate of $\alpha$ less than one, the model clearly is not appropriate. Thus, considerations of invariance imposed by averaging help eliminate a popular model of deviations from Weber's law.

It turns out, though, that not all power function models share such severe mathematical constraints on their parameter values under this averaging. In particular, one model which is not so severely constrained, yet provides a good fit for many intensity discrimination data, presents $\xi_\nu$ as a power function, i.e., as

$$\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}. \tag{7}$$

in which $\beta(\nu) > 0$ and $K(\nu) > 0$ are parameters that may depend upon the value $\nu$ of the criterion. Chapters 4 and 5 contain several empirical and theoretical results pertaining to this model, the most important of which may be that the exponent $\beta(\nu)$ in Eq. (7) varies with the discrimination criterion $\nu$. This result indicates that the value of the near-miss exponent depends on the definition of 'just-noticeable' in the estimation of $x + \Delta(x)$. This should be a caution against regarding the exponent as a critical aspect of neural coding of acoustic intensity, as has been the tendency in the near-miss literature (cf. Falmagne, 1985; see Viemeister, 1972; Schacknow and Raab, 1973; Moore and Raab, 1974; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Florentine, 1986; Viemeister and Bacon, 1988; Schroder et al., 1994; Gallégo and Micheyl, 1998). It also is argued, in Chapter 5, that the near-miss model (7) may be specialized into the submodel

(8) $$\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)},$$

with $x_*$ and $y_*$ parameters that specify the function $K$, that is, $K(\nu) = x_*^{-\beta(\nu)} y_*$. This submodel has an important fixed-point property: for all values of $\nu$, the point $(x, \xi_\nu(x)) = (x_*, y_*)$ satisfies Eq. (8). Furthermore, the empirical estimates of $x_*$ and $y_*$ reported in the study in Chapter 5 are nearly identical for a given listener and condition, with the values corresponding to a large magnitude (105-128 dB SPL). These fixed point estimates and the form of the model specified by Eq. (8) suggest that experimental sound intensities are subjectively evaluated with respect to a high intensity situated at or near the top of the normal range of hearing. This interpretation is consistent with Parker and Schneider (1994), who propose a subjective 'gain control' mechanism which allows the listener to adjust amplification (or attenuation) in the presence of softer (or louder) sounds for improved discriminability (see also Schneider and Parker, 1990). These results, which point to a marriage between quantitative and qualitative descriptions of empirical phenomena, demonstrate the potential power of mathematical modeling, a theme which links the five studies in this dissertation.

The studies are now presented in full detail.

# Chapter 1

# Almost Connected Orders

Individual preferences in opinion polls are often elicited as asymmetric weak orders, i.e., as rankings of the options that allow ties. A respondent's ranking may change from one poll to the next in a more or less systematic way which could involve intermediate states of mind that connect his or her responses in the successive polls. Allowing for the possibility that such states may be less constrained than weak orders, several questions arise. For example, what can be said about relations lying between two weak orders? A specific setting for this question occurs when $W$ and $W'$ are asymmetric weak orders on a set $\mathcal{Y}$, and $W \subset W'$ with no weak order $W''$ that satisfies $W \subset W'' \subset W'$. What kind of relation $R$ satisfies $W \subset R \subset W'$? Any such $R$ must be asymmetric because $W'$ is asymmetric. Less obviously, $R$ must satisfy the semiorder axiom which says that

$$\text{(1.1)} \qquad \text{if } xRy \text{ and } yRz, \quad \text{then } xRw \text{ or } wRz$$

for all $x, y, z, w$ in $\mathcal{Y}$. But $R$ need not be a semiorder (see Theorem 1.15) because it can violate the semiorder axiom which says that if $xRy$ and $zRw$, then $xRw$ or $zRy$.

This paper is devoted to a study of asymmetric relations that satisfy (1.1) and their relationship to weak orders. We refer to such a relation as an 'almost connected order,' or 'ac-order,' because (1.1) is a natural generalization of a connectedness condition (cf. Remark 1.4(b)). Other names have been used for (1.1) and its asymmetric offspring. Chipman (1971) refers to (1.1) as 'semitransitivity,' as do Fishburn (1997) and Fishburn and Trotter (1999). Monjardet (1978) refers to asymmetric relations that satisfy (1.1) as 'S-relations,' and Fishburn (1985) calls them 'partial semiorders.'

The paper recalls previous results and establishes a number of new results for almost connected orders. We prove that every ac-order $R$ is bracketed in a natural way by weak orders $\dot{R}$ and $R^h$ such that $\dot{R}$ is maximum in the set of weak orders included in $R$, and $R^h$ is minimal, but not necessarily minimum, in the set of weak orders that include $R$. We show that the family of ac-orders on a set $\mathcal{Y}$ is not well graded (in the sense of Doignon and Falmagne, 1997) if $|\mathcal{Y}| > 4$. However, we prove that every nonempty ac-order $R$ contains a covering pair $e = (x, y)$ such that $R \setminus \{e\}$ is also an ac-order on the same set. Similarly, we show that, for every ac-order $R$ that is not a chain, there is an $e = (x, y)$ not in $R$ such that $R \cup \{e\}$ is an ac-order.

Some of our results are related to recent work by Trenk (1998) and Gimbel and Trenk (1998). Connections to their work are noted.

## 1.1 Basic Concepts and Preparatory Results

**Definition 1.1.** Except where indicated otherwise, all relations are on a basic finite set $\mathcal{Y}$. Strict and non-strict inclusions are denoted by $\subset$ and $\subseteq$, respectively. The ordered pair $(x, y) \in \mathcal{Y} \times \mathcal{Y}$ is abbreviated as $xy$. A pair $xy$ in a relation $S$ is called a *covering pair* if there is no $z$ in $\mathcal{Y}$ such that $xSzSy$ (which is an abbreviation of '$xSz$ and $zSy$'). In such a case, we may also say that $y$ *covers* $x$. The *product* of two relations $S$ and $T$ is defined by

$$ST = \{xy \mid \exists z \in \mathcal{Y}, xSz \text{ and } zTy\}.$$

We write $S^1 = S$, and for any integer $n \geq 1$, $S^{n+1} = SS^n$. For any relation $S$ on $\mathcal{Y}$, we denote the identity on $\mathcal{Y}$ by $S^0$, and we use $S^{-1}$ to mean the relation $\{yx \mid xSy\}$. The complement $\bar{S}$ of a relation $S$ is in reference to the ground set $\mathcal{Y}$, namely, $\bar{S} = (\mathcal{Y} \times \mathcal{Y}) \setminus S$. We use the formulas $x\bar{S}y$ and $\neg(xSy)$ interchangeably. The *transitive closure* of a relation $S$ is the union

(1.2) $$\tilde{S} = \cup_{n=1}^{\infty} S^n.$$

(Thus, $\tilde{S}$ is not necessarily reflexive.)

**Definition 1.2.** A relation $R$ on $\mathcal{Y}$ is called an *almost connected order* or *ac-order* on $\mathcal{Y}$ if it satisfies the following two properties:

[**AS**] $$R \subseteq \bar{R}^{-1}.$$

[**C2**] $$RR\bar{R}^{-1} \subseteq R.$$

Thus, an ac-order is an asymmetric relation satisfying also [**C2**], which is one of the standard axioms for semiorders (cf. Definition 1.3). Note that [**C2**] is a compact way of writing (1.1). We comment on the '2' in [**C2**] in Remark 1.4(b). Throughout the paper, we write $\mathcal{A}$ for the set of all ac-orders on $\mathcal{Y}$.

Examples and counterexamples for ac-orders are displayed in Figure 1.1 by their Hasse diagrams (see Examples 1.5 below). Note that any asymmetric relation $R$ satisfying $RR = \emptyset$ is an ac-order since Condition [**C2**] holds vacuously. On the other hand, it is easily shown that any ac-order is necessarily irreflexive and transitive (cf. Proposition 1.6 (i) and (ii)), that is, such a relation is a *(strict) partial order* or *poset*.

We also introduce three well-known classes of relations closely related to ac-orders.

**Definition 1.3.** A *(strict) weak order* on a set $\mathcal{Y}$ is a relation $W$ on $\mathcal{Y}$ which is asymmetric (thus, [**AS**] holds) and also satisfies

[**C1**] $$W\bar{W}^{-1} \subseteq W.$$

More explicitly, $W$ is a weak order if for all $x$, $y$ and $z$ in $\mathcal{Y}$,

$$xWy \implies [y\bar{W}x \text{ and } (xWz \text{ or } zWy)].$$

Any weak order is an ac-order (cf. Remark 1.8). Posets are said to be *proper* if they are not weak orders. A *semiorder* is an ac-order satisfying the biorder axiom

**[BI]** $$W\bar{W}^{-1}W \subseteq W$$

(see Luce, 1956; Ducamp and Falmagne, 1969; Fishburn, 1971, 1975; Doignon et al., 1984).

**Remarks 1.4.** (a) It is well known (cf. Krantz et al., 1971; Roberts, 1979) that a relation $W$ on an arbitrary set is a weak order if and only if there exists a mapping $f$ of $(\mathcal{Y}, W)$ into a strict linear order $(\mathcal{X}, \prec)$ such that, for all $x, y \in \mathcal{Y}$, $xWy \Leftrightarrow f(x) \prec f(y)$.

(b) Notice that **[C1]** and **[C2]** are two instances of a class of 'connectedness' conditions, differing by the value $n$ of the exponent in the formula

**[Cn]** $$R^n\bar{R}^{-1} \subseteq R.$$

Conditions **[C1]** and **[C2]** arise when $n = 1$ and $n = 2$, respectively, while **[C0]** means $\bar{R}^{-1} \subseteq R$ (because $R^0$ denotes the identity on $\mathcal{Y}$), that is, $R$ is connected in the usual sense. This explains the term 'almost connected' given to relations satisfying **[C2]**.

(c) Some noteworthy results regarding ac-orders have been obtained. For example, Fishburn (1985) showed that the product of any two ac-orders is a semiorder, and Fishburn and Trotter (1999) pointed out that not only the (order) dimension in the sense of Dushnik and Miller (1941), but also the semiorder dimension and the interval order dimension of finite ac-orders, are unbounded. Skandera (2000) recently gave a characterization of ac-orders in terms of their 'antiadjacency' matrices.

(d) Trenk (1998) and Gimbel and Trenk (1998) present a generalization of weak orders which they call $k$-weak orders. Any ac-order is 1-weak (cf. Trenk, 1998, Theorem 17), but there exist 1-weak orders which are not ac-orders, the simplest being our example (g) in Figure 1.1 (cf. Trenk, 1998, Proposition 14). Trenk (1998, Theorem 19) also gives a characterization of 1-weak, 'totally bounded bitolerance orders,' and we note that some proper ac-orders are totally bounded bitolerance orders, such as the one in Example (c) of Figure 1.1, and others are not, such as the 'standard example of a $n$-dimensional poset' for, say, $n = 3$ (see Trotter, 1992).

**Examples 1.5.** Throughout the paper, we represent the ordered pair $xy$ by an edge going down from $y$ to $x$. (Thus, the relation in Figure 1.1, Example (d) includes the pairs 21, 41, 83, etc.) Example (a) is borrowed from Fishburn (1985). It is a semiorder, a special case of an ac-order (cf. Remark 1.8). Example (b) is a weak order. Examples (c) and (d) are ac-orders which are neither semiorders nor weak orders. Each of Examples (f) and (g) satisfies exactly one of the two axioms [**AS**] and [**C2**] of ac-orders. The failing axiom is indicated at the bottom of each graph. Neither of Examples (e) and (h) is an ac-order.

Since all of the results in the proposition below are either known (see Chipman, 1971; Monjardet, 1978) or immediate, we omit the proof.

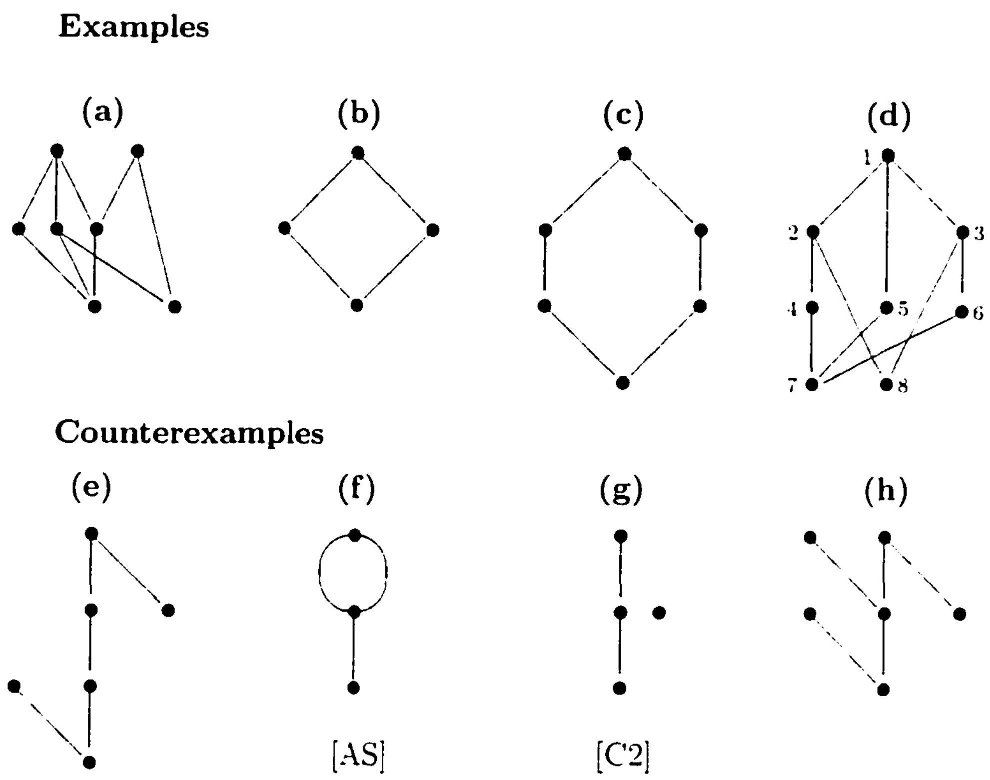


Figure 1.1: Examples and counterexamples of ac-orders represented by their Hasse diagrams.

**Proposition 1.6.** *If $R$ is an ac-order on $\mathcal{Y}$, then*

(i) *$R$ is irreflexive;*

(ii) *$R$ is transitive;*

(iii) $\bar{R}^{-1}RR \subseteq R$;

(iv) $\bar{R}\bar{R}R^{-1} \subseteq \bar{R}$;

(v) $R^{-1}\bar{R}\bar{R} \subseteq \bar{R}$.

**Definition 1.7.** For any weak order $W$ on $\mathcal{Y}$, we write $\sim_W$, or more simply $\sim$ when no ambiguity can arise, for the equivalence relation induced by $W$ on $\mathcal{Y}$: that is, $x \sim_W y$ if and only if $\neg(xWy)$ and $\neg(yWx)$. Any element of the partition of $\mathcal{Y}$ induced by $\sim_W$ is called an *(equivalence) class* of $W$. The particular class of $W$ containing some $x \in \mathcal{Y}$ is denoted by $[x]_W$ (or more simply by $[x]$). If $C$ and $D$ are two classes of $W$, we say that $D$ *covers* $C$ (for $W$) if for all $x \in C$ and $y \in D$, we have $xWy$ and $\neg(xWWy)$.

We apply the concept of covering pair to the inclusion relation for the weak orders in the collection $\mathcal{W}$ of all the weak orders on $\mathcal{Y}$. More precisely, for any two $W, W' \in \mathcal{W}$, we say that $(W, W')$ is a *covering pair* when $W \subset W'$ and there is no $W'' \in \mathcal{W}$ such that $W \subset W'' \subset W'$.

**Remark 1.8.** All linear orders, weak orders, and semiorders are ac-orders, which themselves are posets. Writing $\mathcal{L}$, $\mathcal{S}$ and $\mathcal{P}$, respectively, for the classes of linear orders, semiorders and posets on $\mathcal{Y}$, we have actually

$$\mathcal{L} \subseteq \mathcal{W} \subseteq \mathcal{S} \subseteq \mathcal{A} \subseteq \mathcal{P}, \tag{1.3}$$

with all four inclusions strict if $|\mathcal{Y}| \geq 4$.

As the facts gathered in the next two lemmas are common knowledge, we omit the proofs.

**Lemma 1.9.** *Suppose that $W$ is a weak order on $\mathcal{Y}$.*

(i) *If $C$ is a class of $W$ and $\{C_1, C_2\}$ a partition of $C$, then the relation $S = W \cup (C_1 \times C_2)$ is a weak order. The sets $C_1$ and $C_2$ are classes of $S$, with $C_2$ covering $C_1$ in the sense of Definition 1.7.*

(ii) *Conversely, if both $C_1$ and $C_2$ are classes of $W$, with $C_2$ covering $C_1$, then the relation $T = W \setminus (C_1 \times C_2)$ is a weak order and $C_1 \cup C_2$ is a class of $T$.*

**Lemma 1.10.** *Let $W$ and $W'$ be two weak orders on $\mathcal{Y}$ and suppose that $W \subset W'$. For any $x \in \mathcal{Y}$, we write*

$$[x] = \{y \in \mathcal{Y} \mid y \sim_W x\} \quad \text{and} \quad [x]' = \{y \in \mathcal{Y} \mid y \sim_{W'} x\}.$$

*Then, there exist $s, t$ in $\mathcal{Y}$ such that*

(i) *$s \sim_W t$;*

(ii) *$sW't$;*

(iii) *there is no $z \in \mathcal{Y}$ such that $sW'zW't$.*

*Moreover, if $(W, W')$ is a covering pair (cf. Definition 1.7), then*

(iv) *$\{[s]', [t]'\}$ is a partition of $[s] = [t]$;*

(v) *$W' = W \cup ([s]' \times [t]')$; thus, $[t]'$ covers $[s]'$ for $W'$.*

The next definition and theorem recall a standard concept and apply it in our context.

**Definition 1.11.** For any relation $S$, we denote by $\|_S$ the *incomparability relation* of $S$: thus, for all $x$ and $y$ in $\mathcal{Y}$,

$$x \|_S y \iff (x\bar{S}y \text{ and } y\bar{S}x).$$

When no ambiguity can arise, we abbreviate $\|_S$ as $\|$. We may also write $x \| \{x_1, x_2, \ldots\}$ to mean $x \| x_1, x \| x_2, \ldots$ and use other self explanatory shorthand notation.

We recall from (1.2) that $\tilde{\|}$ denotes the transitive closure of $\|$. It is clear that the transitive closure of the incomparability relation of an irreflexive relation on a set is an equivalence relation on that set. In particular, $\tilde{\|}_P$ is an equivalence relation on $\mathcal{Y}$ for any poset $P$ on $\mathcal{Y}$.

**Theorem 1.12.** *A poset $P$ on $\mathcal{Y}$ is a weak order if and only if $\tilde{\|} = \|$. Thus, a poset $P$ is proper if and only if there exists at least one class $C$ of the partition of $\mathcal{Y}$ induced by $\tilde{\|}$ such that $xPy$ for some $x, y \in C$.*

We omit the proof.

**Definition 1.13.** Let $P$ be a proper poset on $\mathcal{Y}$. Any class $C$ of the partition induced by $\tilde{\|}_P$ satisfying the condition of Theorem 1.12 is called a *critical* class of $P$. A pair $xy$ in $P$ such that $x$ and $y$ belong to the same critical class is called a *critical pair* of $P$.

**Remark 1.14.** (a) By definition, any proper ac-order has at least one critical class. Example (d) in Figure 1.1 displays the Hasse diagram of a proper ac-order on a set of 8 points. The reader can verify that the induced equivalence relation has two classes, $\{1\}$ and $\{2, \ldots, 8\}$. The latter is the only critical class of the ac-order.

(b) Trenk (1998, Proposition 6) describes Algorithm Stackem which acts on a poset and produces a partition of the ground set by identifying 'inseparable' induced suborders. When applied to a poset $P$, this algorithm induces the same partition as does the equivalence relation $\tilde{\|}_P$.

We now turn to the first question raised in our introduction.

## 1.2 What Relation Can Be Squeezed Between Two Weak Orders ?

**Theorem 1.15.** *Let $W, W'$ be two weak orders forming a covering pair. If $W \subset R \subset W'$, then $R$ is an ac-order. Moreover,*

*(i) $R$ has a single critical class $C_1 \cup C_2$, with $C_1 \times C_2 = W' \setminus W$;*

*(ii) $\|_{W'} \subset \|_R \subset \tilde{\|}_R = \|_W$;*

*(iii) $RR \cap \tilde{\|}_R = \emptyset$.*

*Conversely, any ac-order $R$ with a single critical class and such that (iii) holds satisfies $W \subset R \subset W'$ for some covering pair $(W, W')$ of weak orders. Such a pair $(W, W')$ is unique iff any element of the critical class appears in some critical pair of $R$.*

PROOF. Since $W'$ is asymmetric and $R \subset W'$ by hypothesis, $R$ must also be asymmetric. Let $s$ and $t$ be as in Lemma 1.10 and set $C_1 = [s]'$ and $C_2 = [t]'$; thus,

$$W' = W \cup (C_1 \times C_2). \tag{1.4}$$

$$\emptyset \neq R \setminus W \subset C_1 \times C_2. \tag{1.5}$$

We prove that $R$ satisfies **[C2]**. Suppose that

$$xRy, yRz, \text{ and } \neg(wRz). \tag{1.6}$$

We have to establish $xRw$. Since $W \subset R$, (1.6) implies

$$\neg(wWz). \tag{1.7}$$

If both $xWy$ and $yWz$ also hold, then $xWw$ because $W$ is an ac-order, and so $xRw$ because $W \subset R$ by hypothesis. Thus, we suppose that either $\neg(xWy)$ (Case 1), or $\neg(yWz)$ (Case 2).

*Case* 1. If $\neg(xWy)$, then $xRy$ implies $xy \in R \setminus W$, yielding $xy \in C_1 \times C_2$ by (1.5). We cannot have $yz \in C_1 \times C_2$ because $C_1 \cap C_2 = \emptyset$. Thus, $yRz$ leads to $yWz$, which together with (1.7) and the fact that $W$ is a weak order, yields $yWw$. Since $W$ is a weak order, we have either $xWw$ or $yWx$. But $yWx$ together with $xRy$ would give $yW'x$ and $xW'y$, contradicting the asymmetry of $W'$ (as a weak order). We obtain thus $xWw$, and so $xRw$.

*Case* 2. The argument follows the same pattern. Suppose that $\neg(yWz)$. Because $yRz$, this implies $yz \in C_1 \times C_2$, and we cannot have also $xy \in C_1 \times C_2$

because $C_1 \cap C_2 = \emptyset$. Thus $xRy$ implies $xW'y$, which gives $wW'y$ or $xW'w$. But $wW'y$ and $\neg(wW'z)$ (cf. (1.7)) give $zW'y$, which together with $yRz$ would give $zW''y$ and $yW''z$, contradicting the asymmetry of $W''$. We obtain $xW'w$, and so $xRw$, as asserted.

*Proof of* (i). Because the inclusion in (1.5) is strict, there exists $zw \in C_1 \times C_2$ such that $\neg(zRw)$, and also $\neg(wRz)$ (otherwise $zW''w$ and $wW''z$, contradicting the asymmetry of $W''$). Thus, $z \parallel_R w$ holds, with $z \in C_1$ and $w \in C_2$. For any $x, y \in C_i$, with $i = 1$ or $i = 2$, we have $x \parallel_R y$ since each of $C_1$ and $C_2$ is an equivalence class of $W''$ and $R \subset W''$. Also, if $x \in C_1$ and $y \in C_2$, we have $x \parallel_R z \parallel_R w \parallel_R y$. Thus, for any $x, y \in C_1 \cup C_2$ we have $x \overline{\parallel}_R y$, and so $C_1 \cup C_2$ is an equivalence class of $\overline{\parallel}_R$ which is a critical class of $R$ because, by (1.5), we have $R \cap (C_1 \times C_2) \neq \emptyset$. It is in fact the only critical class of $R$. Indeed, suppose that $x \overline{\parallel}_R y$; then, for some sequence $x = x_1, \ldots, x_n = y$, we have $x_i \parallel_R x_{i+1}$, and thus also $x_i \parallel_{W'} x_{i+1}$, $1 \leq i < n$. By the transitivity of $\parallel_{W'}$, this gives $x \parallel_{W'} y$, yielding $\neg(xW'y)$. Using (1.5), we have $xRy$ only if $xy \in C_1 \times C_2$.

*Proof of* (ii). The hypothesis $W' \subset R \subset W''$ implies $\parallel_{W''} \subseteq \parallel_R \subseteq \parallel_{W'}$, and neither of the inclusions can be an equality because $R$ is not a weak order (cf. Theorem 1.12). Turning to the equality $\overline{\parallel}_R = \parallel_{W'}$, suppose that $x \overline{\parallel}_R y$. We must have then either $x \parallel_R y$, and so $x \parallel_{W'} y$ (because $W' \subset R$), or $xy \in \overline{\parallel}_R \cap (R \cup R^{-1})$. This last case subdivides into two subcases. If $xy \in \overline{\parallel}_R \cap R$, then there exists a sequence $x = x_1, \ldots, x_n = y$ such that $x_i \parallel_R x_{i+1}$, and thus $x_i \parallel_{W'} x_{i+1}$, for $1 \leq i < n$. This yields $x \parallel_{W'} y$ by transitivity. The other subcase $xy \in \overline{\parallel}_R \cap R^{-1}$ follows by symmetry. This proves that $\tilde{\|}_R \subseteq \|_W$. Conversely, if $x \|_W y$, then either $x \|_R y$ or $xy \in C_1 \times C_2$: thus $x \tilde{\|}_R y$ in both cases since by (i), $C_1 \cup C_2$ is an equivalence class of $\tilde{\|}_R$.

*Proof of* (iii). Suppose that $x(RR \cap \tilde{\|}_R)y$. Since $R$ is transitive by Proposition 1.6(ii), we get $x(R \cap \tilde{\|}_R)y$, which implies that $x$ and $y$ belong to the single critical class of $R$ (cf. (i) above). Thus, $xy \in C_1 \times C_2$, which together with $xRRy$ contradicts the fact that $C_2$ covers $C_1$ (see Lemma 1.10(v)).

We leave the converse to the reader. □

Next, we consider an arbitrary ac-order $R$ and we ask: what are weak orders $W'$ and $W''$ such that $W' \subset R \subset W''$, with $W'$ maximal (or maximum) and $W''$ minimal (or minimum)? Our results in the next two sections transcend ac-orders and apply in fact to general posets.

## 1.3 Contraction of a Poset

**Definition 1.16.** We define the *(weak order) contraction* of a poset $P$ by the formula

$$\dot{P} = P \setminus \tilde{\|}. \tag{1.8}$$

where $\tilde{\|}$ denotes the transitive closure of the incomparability relation $\|$ of $P$. (We thus abbreviate the notation of $\|_P$ and $\tilde{\|}_P$.)

Note that we have $\breve{P} = P$ if $P$ is a weak order: in such a case, we have $\tilde{\|} = \|$ and $P \cap \tilde{\|} = \emptyset$. When $P$ is not a weak order, however, $\breve{P}$ is a proper subset of $P$. In the next few pages, we investigate the properties of such a contraction.

These concepts apply to ac-orders, which are a special case of posets (cf. Proposition 1.6(i)-(ii) and Remark 1.8). We first give an example of a particular ac-order $R$ and its contraction $\breve{R}$. These relations are represented by their Hasse diagrams in Figure 1.2 (ignore the relation $R^h$ for the moment: cf. Definition 1.19). In this example, the ac-order $R$ has a single critical class $C = \{a, b, c, d\}$, the transitive closure $\tilde{\|}_R$ of the incomparability relation of $R$ has three equivalence classes, and we have

$$\breve{R} = R \setminus (\{c, d\} \times \{a, b\}).$$

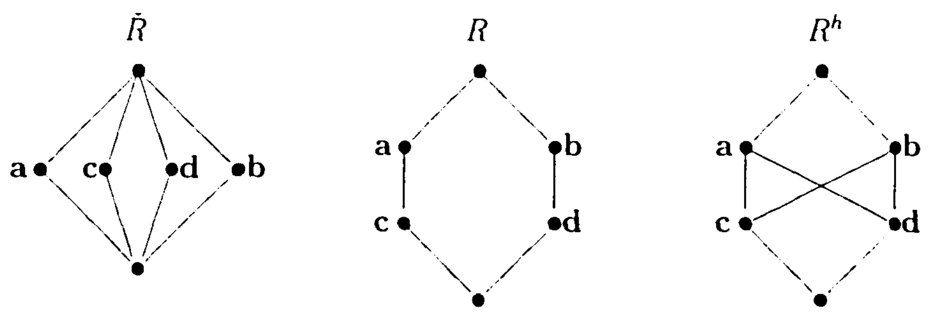


Figure 1.2: An ac-order $R$, its contraction $\breve{R}$, and its height weak order $R^h$, all represented by their Hasse diagrams (cf. Definitions 1.16 and 1.19).

**Theorem 1.17.** *The contraction $\check{P}$ of a poset $P$ is a weak order, which is the maximum of all those weak orders included in $P$. Moreover, $\tilde{\|}_P$ coincides with the equivalence relation induced by the weak order $\check{P}$; hence, all the critical classes of $P$ (if any exist) are equivalence classes of $\check{P}$.*

Obviously, $\check{P}$ may have equivalence classes which are not critical classes of $P$. Also, if $P$ has no critical classes, then $P = \check{P}$ (cf. Theorem 1.12).

PROOF. The definition of $\check{P}$ logically implies $\|_{\check{P}} = \tilde{\|} \cup \| = \tilde{\|}$. Thus, the equivalence relation $\tilde{\|}$ coincides with the incomparability relation of $\check{P}$, and since any critical class of $P$ is an equivalence class of the partition induced by $\tilde{\|}$, it is also an equivalence class of $\|_{\check{P}}$.

We show that $\check{P}$ is a weak order. Note that $\check{P}$ is asymmetric, since $P$ is. Note also that $\check{P}$ is transitive. Indeed, if $x\check{P}y\check{P}z$, then $xPz$ because $\check{P}\check{P} \subseteq PP \subseteq P$. This gives $\neg(z\check{P}x)$ by the asymmetry of $P$. We have thus either $x\check{P}z$ or $x \|_{\check{P}} z$. In the latter case, there exists a sequence $x_0 = x, x_1, \ldots, x_{n+1} = z$ such that $x_0 \| x_1 \| \ldots \| x_{n+1}$. Note that we cannot have $x_1 \|_{\check{P}} y$, because together with $x_0 \| x_1$, this would give $x \|_{\check{P}} y$, contradicting $x\check{P}y$. We have thus either $x_1\check{P}y$ or $y\check{P}x_1$. If the latter holds, then $xPx_1$ by transitivity of $P$, contradicting $x_0 \| x_1$. So, we obtain $x_1\check{P}y$. By a similar argument, we obtain $x_2\check{P}y$, $x_3\check{P}y$, $\ldots$, $x_n\check{P}y$, giving $x_n\check{P}y\check{P}z$. But this leads to $x_nPz$, which contradicts $x_n \| z$. Thus, it is impossible that $x \|_{\check{P}} z$, so $x\check{P}y\check{P}z$ implies $x\check{P}z$.

We now prove that any weak order $W'$ included in $P$ is also included in $\check{P}$. Assume $xW'y$. We have thus also $xPy$. If $xy$ is not a critical pair of $P$, then

$xy \notin \tilde{\|}$, yielding $xy \in \dot{P} = P \setminus \tilde{\|}$. Accordingly, let $xy$ be a critical pair of $P$, that is, $x(P \cap \tilde{\|})y$. This implies that there must be a sequence $x_1 = x, x_2, \ldots, x_n = y$ such that $x_i \| x_{i+1}$ for $1 \leq i < n$. Since $W \subset P$ by hypothesis, we also have $x_i \|_W x_{i+1}$ for $1 \leq i < n$, which implies $x \|_W y$, contradicting our hypothesis that $xWy$. We conclude that any $xy \in W \subset P$ cannot be a critical pair of $P$ and must belong to $\dot{P}$. Consequently, $\dot{P}$ is indeed the maximum of all those weak orders included in $P$. This completes the proof of Theorem 1.17. □

**Remark 1.18.** Defining the *disconnected degree* of a poset $P$ to be the number $\delta(P)$ of its critical classes, we have that any proper poset $P$ can be represented as a union of $\delta(P)$ posets, each of which has only one critical class (and is thus of disconnected degree 1). Indeed, it can be shown that:

*If $P$ is a proper poset and $\mathcal{K}(P)$ is the collection of its critical classes, then for any $C \in \mathcal{K}(P)$ the relation $\dot{P} \cup (P \cap (C \times C))$ is a poset of disconnected degree 1, and*

$$P = \cup_{C \in \mathcal{K}(P)} \left(\dot{P} \cup (P \cap (C \times C))\right). \tag{1.9}$$

## 1.4 Expansions of a Poset

We now discuss minimal weak orders including a poset $P$. The situation is not as clear cut as in the case of the contraction because (as will be proved) a minimal weak order including $P$ is unique only if $P$ itself is this weak order. Candidates for minimal weak orders are easily derived from the 'height' and the 'depth' of $P$.

**Definition 1.19.** Suppose that $P$ is a poset on the finite set $\mathcal{Y}$. The *height* $h(x)$ of a point $x$ in $\mathcal{Y}$ equals the cardinality of a longest chain within $P$ ending in $x$. Thus, $xPy$ implies $h(x) < h(y)$ (but not conversely), and the minimal elements of $P$ have height 1. When the set $H_k = \{x \in \mathcal{Y} \mid h(x) = k\}$ is not empty, it is referred to as the *level* $k$ (of $P$). The *height weak order* $P^h$ of $P$ is defined by $xP^h y$ iff $h(x) < h(y)$ (cf. Remark 1.4(a)). The *height* $h(P)$ of poset $P$ is the largest value of $h$ on $\mathcal{Y}$.

Similarly, the *depth* $d(x)$ of $x$ in $\mathcal{Y}$ is the cardinality of a longest chain starting in $x$. Maximal elements of $P$ have depth 1. The *depth weak order* $P^d$ is defined by $xP^d y$ iff $d(x) > d(y)$. Clearly, both $P^h$ and $P^d$ are weak orders including $P$.

**Remark 1.20.** The height and depth weak orders coincide iff all maximal chains in $P$ have the same length (here, 'maximal' means 'maximal for inclusion'). In particular, $P^h = P^d = P$ if $P$ is a weak order.

**Lemma 1.21.** *Both $P^h$ and $P^d$ are minimal among those weak orders including $P$.*

PROOF. Suppose $P \subseteq W \subseteq P^h$ for some weak order $W$. Take a longest chain $x_1 P x_2 P \cdots P x_h$ in $P$. All these elements must belong to distinct classes of $W$, and also of $P^h$. Pick $y \in \mathcal{Y}$. Then we cannot have $y W x_1$ (because $W \subseteq P^h$), and neither $x_n W y$. Assume now $x_i W y W x_{i+1}$ for some $i = 1, 2, \ldots, h$. As $W \subseteq P^h$, we get $h(x_i) < h(y) < h(x_{i+1})$, which contradicts the maximality of the chosen chain. The argument for $P^d$ is similar. □

**Example 1.22.** With $R$ the ac-order whose Hasse diagram is given in Figure 1.2, we have that $R^h = R^d$. Note that $R \cup \{dc, da, ba\}$ is another weak order that minimally includes $R$.

**Theorem 1.23.** *Any proper poset is included in at least two minimal weak orders.*

PROOF. Let $P$ be a proper poset. By Remark 1.20 and Lemma 1.21, if $P$ has two maximal chains with different length, then $P^h$ and $P^d$ may be taken to be those two weak orders. Assume thus that all maximal chains of $P$ have same length. Since $P$ is not a weak order, there exist $x$, $y$, $z$ in $\mathcal{Y}$ such that $xPy$ and $z \parallel \{x, y\}$. Clearly, we may assume that $y$ covers $x$. Take any maximal chain $C$ of $P$ containing $z$. Then there are $u$ and $v$ in $C$ with

$$u \neq x, \quad v \neq y, \quad h(u) = h(x), \quad h(x) + 1 = h(v) = h(y),$$

$$\text{and either} \quad zPv \text{ (Case 1)} \quad \text{or} \quad uPz \text{ (Case 2)}.$$

In Case 1, we necessarily have $uy \in P^h \setminus P$ and it is easily seen that

$$(P^h \setminus \{uy\}) \cup \{wu \mid h(w) = h(x) \text{ and } w \neq u\} \cup \{yt \mid h(t) = h(v) \text{ and } t \neq y\}$$

is a minimal weak order including $P$ and distinct from $P^h$. A similar construction can be performed in Case 2. □

**Remark 1.24.** A representation of a proper ac-order $R$ similar to that in Remark 1.18 can be formulated in terms of $R^h$, involving an intersection of $\delta(R)$ ac-orders of disconnected degree 1.

## 1.5 Non-Wellgradedness and the Fringes of an AC-Order

**Definition 1.25.** Let $S$ be any set in a family $\mathcal{F}$ of subsets of some set $\mathcal{X}$, not necessarily finite. The *outer fringe* of $S$ (with respect to $\mathcal{F}$) is the set $S^{\mathcal{O}}$ containing all the points $x \in \bar{S} = \mathcal{X} \setminus S$ such that $S \cup \{x\}$ is another set in the family $\mathcal{F}$; formally,

$$S^{\mathcal{O}} = \{x \in \bar{S} \mid S \cup \{x\} \in \mathcal{F}\}.$$

Similarly, the *inner fringe* of $S$ (with respect to $\mathcal{F}$) is the set $S^{\mathcal{I}}$ containing all those points $x \in S$ such that $S \setminus \{x\}$ is another set in $\mathcal{F}$; formally,

$$S^{\mathcal{I}} = \{x \in S \mid S \setminus \{x\} \in \mathcal{F}\}.$$

For example, the outer fringe of an ac-order $R$ in $\mathcal{A}$ is the relation $R^{\mathcal{O}}$ containing all those pairs $xy$ in $\bar{R}$ such that $R \cup \{xy\} \in \mathcal{A}$. The concepts of inner and outer fringes of a set in a family of subsets were introduced by Doignon and Falmagne (1997); see also Doignon and Falmagne (1999). We specify the inner and outer fringes of an ac-order $R$ directly in terms of $R$, as a step in a proof that the family of all ac-orders on a finite set is not, in general, 'well graded' (see Definition 1.27 and Theorem 1.29).

**Proposition 1.26.** *The inner and outer fringes $R^{\mathcal{I}}$ and $R^{\mathcal{O}}$ of an ac-order $R$ on $\mathcal{Y}$, with respect to the family $\mathcal{A}$ of all ac-orders on $\mathcal{Y}$, are respectively given by*

$$R^{\mathcal{I}} = R \setminus (RR\bar{R}^{-1} \cup \bar{R}^{-1}RR), \tag{1.10}$$

$$R^{\mathcal{O}} = \bar{R} \setminus (R^{0} \cup \bar{R}\bar{R}R^{-1} \cup R^{-1}\bar{R}\bar{R}). \tag{1.11}$$

PROOF. Let $R$ be an ac-order. We first prove (1.10). Supposing $xR^{\mathcal{I}}y$, we have by definition of $R^{\mathcal{I}}$ that $xRy$ and that $R \setminus \{xy\}$ is an ac-order. If $xRR\bar{R}^{-1}y$, then there exist $x_1, x_2 \in \mathcal{Y}$ such that $xRx_1$, $x_1Rx_2$, and $\neg(yRx_2)$. Note that $x_1 \neq x$ (by irreflexivity of $R$), $x_2 \neq x$ (by irreflexivity and transitivity of $R$), and $x_1 \neq y$ (since then $yRx_2$ and $\neg(yRx_2)$). Furthermore, it can not be that $x_2 = y$ with $x_1 \neq x$ and $x_1 \neq y$, since $R \setminus \{xy\}$ is transitive. Thus $xRx_1Rx_2\bar{R}^{-1}y$, with neither $x_1$ nor $x_2$ in $\{x, y\}$. Since this clearly violates [**C2**] in $R \setminus \{xy\}$, it must be that $xy \notin RR\bar{R}^{-1}$. Similarly, $xy \notin \bar{R}^{-1}RR$, so $R^{\mathcal{I}} \subseteq R \setminus (RR\bar{R}^{-1} \cup \bar{R}^{-1}RR)$.

For the reverse inclusion in (1.10), suppose $wz \in R \setminus (RR\bar{R}^{-1} \cup \bar{R}^{-1}RR)$. We wish to show that $R \setminus \{wz\}$ is an ac-order. Note that $R \setminus \{wz\}$ is asymmetric since $R$ is. Suppose $R \setminus \{wz\}$ did not satisfy [**C2**]. Because $R$ is an ac-order, with $R$ and $R \setminus \{wz\}$ differing only by the element $wz$, we would then have the existence of $w_1, w_2 \in \mathcal{Y}$ such that $w_1Rz$, $w_2Rw_1$, and $\neg(w_2Rw)$, or we would have the existence of $z_1, z_2 \in \mathcal{Y}$ such that $z_2Rz_1$, $wRz_2$, $and \neg(zRz_1)$. But both of these are impossible, the former implying $w\bar{R}^{-1}RRz$, and the latter implying $wRR\bar{R}^{-1}z$. Thus, the assumption that $R \setminus \{wz\}$ does not satisfy [**C2**] is false, so $R \setminus \{wz\}$ is an ac-order, and the equality in (1.10) is established.

For (1.11), suppose first that $xR^{\mathcal{O}}y$. Then $x\bar{R}y$, and we must show $\neg(xR^0y)$, $xy \notin \bar{R}\bar{R}R^{-1}$, and $xy \notin R^{-1}\bar{R}\bar{R}$. Clearly we have $\neg(xR^0y)$, for $R \cup \{xy\}$ is an ac-order (and hence irreflexive) by definition of $R^{\mathcal{O}}$. If $x\bar{R}\bar{R}R^{-1}y$, there would exist $x_1, x_2 \in \mathcal{Y}$ such that $\neg(xRx_1)$, $\neg(x_1Rx_2)$, and $yRx_2$. Since $yRx_2$ and $R \cup \{xy\}$ is an ac-order, this contradicts [**C2**]. Thus, $xy \notin \bar{R}\bar{R}R^{-1}$. Supposing $xR^{-1}\bar{R}\bar{R}y$, we would have the existence of $y_1, y_2 \in \mathcal{Y}$ such that $y_1Rx$, $\neg(y_1Ry_2)$, and $\neg(y_2Ry)$. Since $xRy$ and [**C2**] holds for $R \cup \{xy\}$, we get a contradiction. This establishes the inclusion $R^{\mathcal{O}} \subseteq \bar{R} \setminus (R^0 \cup \bar{R}\bar{R}R^{-1} \cup R^{-1}\bar{R}\bar{R})$.

Suppose now that $wz \in \bar{R} \setminus (R^0 \cup \bar{R}\bar{R}R^{-1} \cup R^{-1}\bar{R}\bar{R})$. We must show $R \cup \{wz\}$ is an ac-order. If $R \cup \{wz\}$ were not asymmetric, the asymmetry of $R$ would imply $zRw$. Since $\neg(wR^0z)$ and so $w \neq z$ by hypothesis, we would then have $w\bar{R}w$ (by the irreflexivity of $R$) and $wR^{-1}z$, which together would imply $w\bar{R}\bar{R}R^{-1}z$. This is a contradiction, so $R \cup \{wz\}$ must be asymmetric. It remains to show that $R \cup \{wz\}$ satisfies condition [**C2**]. This condition holds in $R$, but we must check that the addition of the pair $wz$ does not cause [**C2**] to fail for $R \cup \{wz\}$. This could happen in only two cases:

*Case* 1. There exist $w_1, w_2 \in \mathcal{Y}$ such that $w_1 R w R z$, but $\neg(w_2 R z)$ and $\neg(w_1 R w_2)$. In this case, we would have $w R^{-1} w_1 \bar{R} w_2 \bar{R} z$, so $wz \in R^{-1}\bar{R}\bar{R}$, a contradiction.

*Case* 2. There exist $w_1, w_2 \in \mathcal{Y}$ such that $w R z R w_1$, but $\neg(w R w_2)$ and $\neg(w_2 R w_1)$. In this case, we would have $w \bar{R} w_2 \bar{R} w_1 R^{-1} z$, so $wz \in \bar{R}\bar{R}R^{-1}$, a contradiction.

Thus $R \cup \{wz\}$ is an ac-order, giving $R^{\mathcal{O}} \supseteq \bar{R} \setminus (R^0 \cup \bar{R}\bar{R}R^{-1} \cup R^{-1}\bar{R}\bar{R})$ and completing the proof of (1.11). □

**Definition 1.27.** A family $\mathcal{F}$ of subsets of a set $\mathcal{X}$ is *1-connected* if, for any $A, B \in \mathcal{F}$ there exists a finite sequence of sets $A = F_0, F_1, \ldots, F_k = B$ in $\mathcal{F}$ such that $|F_{i-1} \triangle F_i| = 1$, $i = 1, \ldots, k$ (where $\triangle$ stands for the symmetric difference between sets). The family $\mathcal{F}$ is said to be *well graded* if, in addition, we can always make $k = |A \triangle B|$.

This definition applies obviously to family of relations. The following result appears in Doignon and Falmagne (1997).

**Theorem 1.28.** *Assume the set $\mathcal{X}$ is finite. The three following conditions on a family $\mathcal{F}$ of subsets of $\mathcal{X}$ are equivalent:*

(i) *$\mathcal{F}$ is well graded;*

(ii) *any two sets $R$ and $S$ in $\mathcal{F}$ which satisfy $R^{\mathcal{I}} \subseteq S$ and $R^{\mathcal{O}} \subseteq \bar{S}$ must be equal;*

(iii) *any two sets $R$ and $S$ in $\mathcal{F}$ which satisfy $R^{\mathcal{I}} \subseteq S$, $R^{\mathcal{O}} \subseteq \bar{S}$, $S^{\mathcal{I}} \subseteq R$, $S^{\mathcal{O}} \subseteq \bar{R}$ must be equal.*

Using this result, Doignon and Falmagne (1997) showed that the respective families of all partial orders, biorders, interval orders, and semiorders on a given finite set are well graded (for partial orders, this was proved before by Ovchinnikov, 1973; see also Ovchinnikov, 1983). We use the same result to prove the following.

**Theorem 1.29.** *The family $\mathcal{A}$ of all ac-orders on $\mathcal{Y}$ is well graded if and only if $|\mathcal{Y}| \leq 4$.*

PROOF. For $|\mathcal{Y}| \leq 3$, the family $\mathcal{A}$ of all ac-orders on $\mathcal{Y}$ corresponds to the family $\mathcal{P}$ of all partial orders on $\mathcal{Y}$, which is well graded. Suppose that $|\mathcal{Y}| = 4$. We use the equivalence of (i) and (iii) in Theorem 1.28 to show that $\mathcal{A}$ is well graded. To this end, suppose $R, S \in \mathcal{A}$ satisfy $R^{\mathcal{I}} \subseteq S$, $R^{\mathcal{O}} \subseteq \bar{S}$, $S^{\mathcal{I}} \subseteq R$, $S^{\mathcal{O}} \subseteq \bar{R}$. In the case $R^{\mathcal{I}} = R$ and $S^{\mathcal{I}} = S$, we have $S^{\mathcal{I}} \subseteq R = R^{\mathcal{I}} \subseteq S = S^{\mathcal{I}}$, so $R = S$. Otherwise, we have without loss of generality $R^{\mathcal{I}} \subset R$. There are only two possibilities for $R$, namely

$$w \parallel xRyRzR^{-1}w \text{ (Case 1)} \qquad \text{and} \qquad wR^{-1}xRyRz \parallel w \text{ (Case 2)}.$$

In Case 1, note that $R^{\mathcal{I}} = \{yz, xy\}$, and since by hypothesis $R^{\mathcal{I}} \subseteq S$, we must have $xSySz$. Also, because $R^{\mathcal{O}} = \{wy, xw\}$ and by hypothesis $R^{\mathcal{O}} \subseteq \bar{S}$, it must be that $w\bar{S}y$ and $x\bar{S}w$. But $w\bar{S}y$ and $x\bar{S}w$ respectively imply (along with the transitivity of $S$) $w\bar{S}x$ and $y\bar{S}w$. Since $S$ satisfies [**C2**], we have necessarily $wSz$, i.e., $S = R$. The proof that we also have $S = R$ in Case 2 is similar. By Theorem 1.28, then, $\mathcal{A}$ is well graded if $|\mathcal{Y}| = 4$.

Figure 1.3 shows that $\mathcal{A}$ is not well graded if $|\mathcal{Y}| = n \geq 5$. In this figure, $R_n^{\mathcal{I}} = \{31, 42\} = S_n^{\mathcal{I}}$ and $R_n^{\mathcal{O}} = \{32, 41\} = S_n^{\mathcal{O}}$, yet $R_n$ and $S_n$ are distinct. As (iii) of Theorem 1.28 is not satisfied, $\mathcal{A}$ is not well graded. □

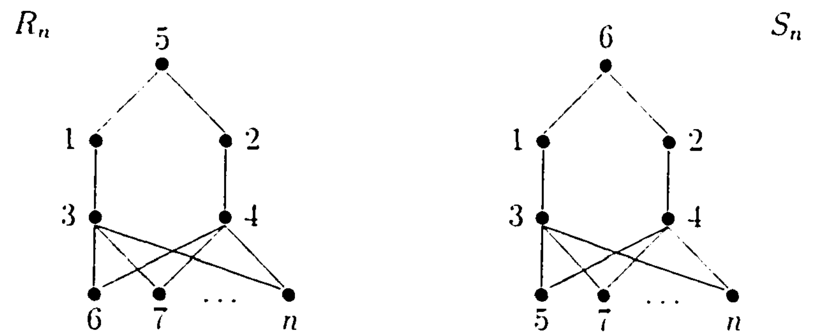


Figure 1.3: Distinct ac-orders on an $n$-element set, $n \geq 5$, which have identical inner and outer fringes. (When $n = j$, the vertices greater than $j$ are omitted.)

## 1.6 Downgradability and Upgradability

We just proved that the family of ac-orders on $\mathcal{Y}$ was not necessarily well graded. However, such a family is always 1-connected (cf. Definition 1.27). In fact, much more can be proved, namely, any nonempty ac-order on $\mathcal{Y}$ can be 'trimmed down' by removing pairs one by one, until the empty ac-order is reached, without ever leaving the family of all ac-orders on $\mathcal{Y}$. Also, going upward, any ac-order which is not a linear order can be enlarged by adding pairs one by one, until a linear order is formed, without ever leaving the family. This section is devoted to the relevant definitions and exact results.

**Definition 1.30.** Any set in a family of sets $\mathcal{F}$ is called *downgradable* (resp. *upgradable*) if it has a nonempty inner fringe (resp. outer fringe). The family $\mathcal{F}$ itself is *downgradable* if all its non-minimal sets are downgradable. It is said to be *upgradable* if all but its maximal sets are upgradable.

**Remarks 1.31.** (a) By itself, neither downgradability nor upgradability implies 1-connectedness. In the case of ac-orders, however, the empty relation is trivially an ac-order and so downgradability (see Theorem 1.32) implies 1-connectedness.

(b) Clearly, if a family of sets is well graded, it is also upgradable and downgradable.

(c) Upgradability and downgradability of a set in a family $\mathcal{F}$ are relative to that family. However, if $\mathcal{F} \subseteq \mathcal{G}$ are two families of sets, and $S \in \mathcal{F}$ is downgradable (resp. upgradable) in $\mathcal{F}$, then is is also downgradable (resp. upgradable) in

$\mathcal{G}$. In particular, any semiorder on $\mathcal{Y}$ which is neither empty nor a linear order is both upgradable and downgradable in $\mathcal{A}$.

**Theorem 1.32.** *The family $\mathcal{A}$ of all ac-orders on $\mathcal{Y}$ is both downgradable and upgradable.*

Note that the maximal sets in $\mathcal{A}$ are the linear orders. Thus, all the other ac-orders are upgradable.

PROOF. We write $\| = \|_R$. Let $R$ be an ac-order on $\mathcal{Y}$ which is neither empty nor a linear order. By Remark 1.31 (c), if $R$ is a semiorder, then it is both downgradable and upgradable in $\mathcal{A}$. Suppose thus that $R$ is not a semiorder, that is, there exist $x, y, z$ and $w$ in $\mathcal{Y}$ such that $xRy$, $zRw$, but neither $xRw$ nor $zRy$. We claim that both of the following must hold:

[D] $xy \in R^{\mathcal{I}}$;

[U] $xw \in R^{\mathcal{O}}$.

Beginning with Claim [D], we proceed by contradiction and suppose that $xy \notin R^{\mathcal{I}}$; thus, $R \setminus \{xy\}$ is not an ac-order. Since $R \setminus \{xy\}$ is asymmetric (because $R$ is asymmetric) and is not an ac-order, it cannot satisfy [**C2**]. We have thus at least one of the following two cases.

*Case* [D1]. There exist $t \neq x$ and $u$ in $\mathcal{Y}$ with $tRuRy$ and $x \parallel t$. It is easily checked that $x$, $z$, $y$, $w$, $t$, and $u$ are all distinct. Also, $z \parallel y$ and $tRRy$ imply $tRz$ by [**C2**], and we get $tRRw$ with $x \parallel \{t, w\}$, an impossibility because $R$ is an ac-order.

*Case* [D2]. There exist $s \neq y$ and $v$ in $\mathcal{Y}$ with $xRvRs$ and $y \parallel s$. All of $x$, $z$, $y$, $w$, $v$, and $s$ are distinct. But then $x \parallel w$ and $xRRs$ imply $wRs$ by [**C2**], and we have $zRRs$ with $y \parallel \{z, s\}$, again an impossibility.

Turning to Claim [U], assume now that $xw \notin R^O$; thus, $R \cup \{xw\}$ is not an ac-order. We cannot have $wRx$ since it would yield $zRwRxRy$, and so $zRy$ by the transitivity of $R$, contradicting $\neg(zRy)$. This implies that $R \cup \{xw\}$ is asymmetric. As it is not an ac-order, it cannot satisfy [**C2**]. Again, we have only two possible cases.

*Case* [U1]. There exist $t$, $u$ in $\mathcal{Y}$ with $tRx$ and $u \parallel \{t, w\}$, $u \notin \{t, w\}$. We see that $x$, $z$, $y$, $w$, $t$, and $u$ must be distinct. Moreover, $z \parallel y$ and $tRRy$ imply $tRz$, and we have $tRRw$ with $u \parallel \{t, w\}$, in contradiction with $R$ being an ac-order.

*Case* [U2]. There exist $v$, $s$ in $\mathcal{Y}$ with $wRs$ and $v \parallel \{x, s\}$, $v \notin \{x, s\}$. All of $x$, $z$, $y$, $w$, $v$, and $s$ are distinct. But $zRwRs$ together with $z \parallel y$ imply $yRs$; so $xRRs$ and $v \parallel \{x, s\}$, again a contradiction of [**C2**].

We conclude that both claims [D] and [U] are true, and so $R$ is both downgradable and upgradable, which establishes the theorem. □

**Remarks 1.33.** (a) The first part of our proof relies on the previously established result that the family $\mathcal{S}$ of semiorders on $\mathcal{Y}$ is well graded (Doignon and Falmagne, 1997). A more direct proof of the downgradability of $\mathcal{S}$ reveals that a nonempty ac-order always contains a pair connecting level 1 and level 2, or a pair connecting level 2 and level 3, that can be removed to give another ac-order. A more direct proof of the upgradability of $\mathcal{S}$ reveals that, for an ac-order $R$ which is not a chain, there must exist two elements within level $1^*$ of $R$, or between levels $2^*$ and $3^*$ of $R$ (where the * means that successive minimum elements have been discarded before the level is determined) that can be added as a pair to $R$ to give another ac-order.

(b) In view of strengthening Theorem 1.32, we could ask whether, for two given ac-orders $R$ and $S$ with $R \subset S$, we always have $S^{\mathcal{I}} \setminus R \neq \emptyset$ (i.e., whether it is always possible to move from a given ac-order $S$ to another one $R$ included in $S$, deleting one pair at a time). The answer is negative, as seen by taking $S$ as in Figure 1.3 and $R = \{31, 42\}$.

(c) Notice that Theorem 1.32 does not extend to the infinite setting. Indeed, take $\mathcal{Y}$ to be the set $\mathbb{R}$ of all reals. Its usual strict linear ordering $<$ is an ac-order which is not downgradable. Moreover, setting $xPy$ exactly when $x + 1 \leq y$, we get an ac-order $P$ on $\mathbb{R}$ which is not upgradable although it is not a maximal order.

# Chapter 2

# Toward a Graded Generalization of Partial Orders

The aim of this work is a better understanding of families of sets closed under union—which generalize 'partial orders'—via the examination of a progression of types of such families. Motivation for this work is detailed in Chapter 0.

We begin with a nonempty, finite set $Q$ and a family $\mathcal{K}$ of subsets of $Q$. The elements of $Q$ are called *items*. Members $K$ of $\mathcal{K}$ are called *states*. If the family $\mathcal{K}$ contains both $\emptyset$ and $Q$, is closed under union, and has the property that

$$(\nexists\, a, b \in Q)(\forall K \in \mathcal{K} : a \in K \Leftrightarrow b \in K)\ ,$$

then $\mathcal{K}$ is called a *knowledge space* on $Q$. If, in addition, $\mathcal{K}$ is closed under intersection, then it is called a *partially ordinal space*.

The reason for this latter terminology is that there exists a one-to-one correspondence between the collection of all partially ordinal spaces $\mathcal{K}$ on $Q$ and the collection of all 'partial orders' $\mathcal{Q}$ on $Q$. (A *partial order* is a reflexive, transitive, and antisymmetric relation.) This is a classical result due to Birkhoff (1937). (See also Doignon and Falmagne, 1999, Theorem 1.49.) The correspondence is defined through the equivalences

$$pQq \iff (\forall K \in \mathcal{K} : q \in K \Rightarrow p \in K) \tag{2.1}$$

and

$$K \in \mathcal{K} \iff (\forall (p, q) \in \mathcal{Q} : q \in K \Rightarrow p \in K). \tag{2.2}$$

An interpretation of these equivalences is that a pair $(p, q)$ is in $\mathcal{Q}$ precisely when $p$ is a 'prerequisite' item for $q$. We formally define the concept of a 'prerequisite' as follows: $p$ is a *prerequisite* for $q$ in $\mathcal{K}$ if, for all $K \in \mathcal{K}$, we have that $q \in K$ implies $p \in K$. A *background* for an item $q$ is a minimal state of $\mathcal{K}$ containing $q$, with minimality being with respect to set inclusion.

Each item in a partially ordinal space has a unique background. Indeed, if item $q$ had distinct backgrounds $C_1 \in \mathcal{K}$ and $C_2 \in \mathcal{K}$, then $C_1 \cap C_2 \in \mathcal{K}$ would be a set which contains $q$ and which is strictly included in $C_1$, contradicting the minimality of $C_1$. Similarly, it is straightforward to show that a knowledge space on $Q$ is closed under intersection if each item in $Q$ has a unique background. (See Theorem 1.40 of Doignon and Falmagne, 1999.)

Thus, a knowledge space not closed under intersection necessarily has at least one item with more than one background. We introduce (following Doignon and Falmagne, 1999) the concept of a 'surmise function,' which is a function that associates to each item its set of backgrounds.

**Definition 2.1.** A function $\sigma$ from $Q$ to $2^{2^Q}$ is called a *surmise function* on $Q$ if it satisfies the following three conditions for all $q, q' \in Q$ and $C, C' \subseteq Q$:

(1) if $C \in \sigma(q)$, then $q \in C$;

(2) if $q' \in C \in \sigma(q)$, then $C' \subseteq C$ for some $C' \in \sigma(q')$;

(3) if $C, C'' \in \sigma(q)$ and $C'' \subseteq C$, then $C = C'$.

Thus, a knowledge space $\mathcal{K}$ on $Q$ is closed under intersection if, and only if, $|\sigma(q)| = 1$ for all $q \in Q$.

**Definitions 2.2.** Let $\mathcal{K}$ be a knowledge space on $Q$, with $|Q| = n$. For each $j \in \{1, \ldots, n\}$, let $S_j = \{A \in 2^Q \mid 1 \leq |A| \leq j\}$. If $\mathcal{P}_j$ is the relation from $S_j$ to $Q$ defined by the equivalence

$$A\mathcal{P}_j q \iff (\forall K \in \mathcal{K} : A \cap K = \emptyset \Rightarrow q \notin K), \tag{2.3}$$

then $\mathcal{P}_j$ is called the *j-entailment* of $\mathcal{K}$. The family $\mathcal{K}_j$ defined by the equivalence

$$K \in \mathcal{K}_j \iff (\forall (A, q) \in \mathcal{P}_j : A \cap K = \emptyset \Rightarrow q \notin K) \tag{2.4}$$

is called the *space generated by* $\mathcal{P}_j$. If $\mathcal{K}_j = \mathcal{K}$, then the $j$-entailment of $\mathcal{K}$ is said to *recover* $\mathcal{K}$, or, equivalently, that $\mathcal{K}$ is *recovered* by the $j$-entailment of $\mathcal{K}$.

**Remarks 2.3.** (i) The terminology "space generated by" in the above definitions is justified since any family defined via (2.4) is a knowledge space. (ii) With $\mathcal{K}_j$ the space generated by the $j$-entailment of $\mathcal{K}$, we have that $\mathcal{K}_j \subseteq \mathcal{K}$ for each $j$, $j = 1, \ldots, |Q|$. (iii) It turns out that a knowledge space $\mathcal{K}$ is closed under intersection precisely when it is recovered by its 1-entailment.

**Example 2.4.** Let $Q = \{a, b, c, d\}$, and let

$$\mathcal{L} = \{\emptyset, \{a\}, \{d\}, \{a, b\}, \{a, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, Q\}.$$

The 1-entailment $\mathcal{P}_1$ of $\mathcal{L}$ is given by

$$\mathcal{P}_1 = \{(a, a), (a, b), (a, c), (b, b), (c, c), (d, d)\},$$

and the 2 entailment $\mathcal{P}_2$ of $\mathcal{L}$ is given by

$$\mathcal{P}_1 \cup \{(\{a, b\}, a), (\{a, c\}, a), (\{a, d\}, a), (\{a, b\}, b), (\{a, c\}, b), (\{a, d\}, b), (\{b, c\}, b),$$
$$(\{b, d\}, b), (\{a, b\}, c), (\{a, c\}, c), (\{a, d\}, c), (\{b, c\}, c), (\{b, d\}, c), (\{c, d\}, c),$$
$$(\{a, d\}, d), (\{b, d\}, d), (\{c, d\}, d)\}.$$

These entailments generate the respective spaces

$$\mathcal{L}_1 = \{\emptyset, \{a\}, \{d\}, \{a, b\}, \{a, c\}, \{a, d\}, \{a, b, c\}, \{a, b, d\}, \{a, c, d\}, Q\}$$

and

$$\mathcal{L}_2 = \mathcal{L},$$

so that the 2-entailment recovers $\mathcal{L}$, but the 1-entailment does not.

For the space $\mathcal{L}$ above, the item $c \in Q$ has two backgrounds, namely, $\{a, b, c\}$ and $\{a, c, d\}$. Since the space is recovered by a 2-entailment and not by a 1-entailment, and in view of Remark 2.3 (iii), one may ask whether there is a relationship between the number of backgrounds for items in a knowledge space and the smallest integer $j$ such that a $j$-entailment recovers the space. As shown in Theorem 2.7 below, a $k$-entailment will always recover a space whose items have at most $k$ backgrounds. However, there exist knowledge spaces with items having more than $k$ backgrounds which may be recovered by a $k$-entailment. The following example gives such a space.

**Example 2.5.** Let $\mathcal{M}$ be the space generated by the family

$$\mathcal{F} = \{\emptyset, \{a\}, \{b\}, \{b, c\}, \{c, e\}, \{b, d, f\}, \{b, c, d\}, \{c, d, e\}, \{a, d, f\}, \{e\}, \{a, f\}, \{b, f\}\},$$

that is, let $\mathcal{M}$ be the union closure of $\mathcal{F}$. Then $\mathcal{M}$ is a knowledge space on $\{a, b, c, d, e, f\}$ that is recovered by a 2-entailment.[1] Note that the item $d$ in $Q$ has four backgrounds.

**Remark 2.6.** It appears possible to construct, for each $k > 2$ and each $i \in \{2, \ldots, k\}$, an example of a knowledge space which is recoverable by an $i$-entailment but which contains an item (or items) with $k$ backgrounds.

However, we have the following theorem:

[1]This was obtained using Mathematica.

**Theorem 2.7.** Suppose that $\mathcal{K}$ is a knowledge space on $Q$. Suppose also that the maximum number of backgrounds in $\mathcal{K}$ for any item of $Q$ is $\ell$. Then $\mathcal{K}$ may be recovered by an $\ell$-entailment.

PROOF. Let $|Q| = n$. By Theorem 5.5 of Doignon and Falmagne (1999), we have that $\mathcal{K}$ may be recovered by an $n$-entailment. Let $\sigma$ be the surmise function on $Q$. Denoting $\max\{|\sigma(q)| : q \in Q\}$ by $\ell$, we wish to show that $\mathcal{K}$ may be recovered by an $\ell$-entailment $\mathcal{P}_\ell$.

Since we necessarily have that the space $\mathcal{K}_\ell$ generated by $\mathcal{P}_\ell$ includes the space $\mathcal{K}$, we must show that $K \notin \mathcal{K}$ implies $K \notin \mathcal{K}_\ell$. So, suppose $K \notin \mathcal{K}$. By (2.4), there exists $(A, p) \in \mathcal{P}_n$ such that $A \cap K = \emptyset$ and $p \in K$. Let $\sigma(p) = \{\sigma(p)_1, \ldots, \sigma(p)_m\}$, and note that $m \leq \ell$. By (2.3), for each $\sigma(p)_i \in \sigma(p)$ there exists $a(i) \in A$ such that $a(i) \in \sigma(p)_i$. (For, otherwise, $A \cap \sigma(p)_i = \emptyset$ but $p \in \sigma(p)_i$.) Writing $A' = \{a(1), a(2), \ldots, a(m)\}$, we have that $A' \cap K = \emptyset$ (since $A' \subseteq A$) and $|A'| \leq m \leq \ell$. Thus, $K \notin \mathcal{K}_\ell$. □

# Chapter 3

# On Invariance Properties of Empirical Laws

Notions of invariance have played a central role in the investigation of statements considered suitable to be scientific laws. For instance, the classical concept of 'dimensional invariance' has been widely used, via the method of dimensional analysis, in the search for lawful numerical relations among physical variables. The method of dimensional analysis may be employed, for example, in the derivation of the functional description of the motion of a simple pendulum (see e.g. Krantz et al., 1971; Narens, 2002). A related invariance notion, 'meaningfulness,' has been used in the theoretical sciences for seemingly the same purpose as dimensional analysis: scientists seek to describe empirical relationships among variables via functional laws, and putative invariances of the measurement theories of these variables may greatly constrain the possible forms of such laws. The specific use of these and related notions of invariance in the formulation of lawful functional relations may be found, for example, in Luce (1959, 1964, 1990); Luce et al. (1990); Osborne (1970); Falmagne and Narens (1983); Aczél et al. (1986); Kim (1990). The focus of the present paper is a comparison of these two notions of invariance, which are appropriately formalized here in the spirit of Falmagne and Narens (1983). Our main result, which gives insight into the relationship between the two formulations, generalizes a result by these authors. In preparation for a formal presentation, we inspect an example.

The pressure ($P$), volume ($v$), temperature ($t$) and quantity ($n$) of an "ideal" gas are related by the equation

$$P(v, t, n) = R\frac{1}{v}nt, \tag{3.1}$$

in which $R$ is a dimensional constant. Note that the numerical value of $R$ depends on the units employed in the measurement of the variables. Let us fix some triple of units in Eq. (3.1), say, liters, kelvin, and moles. Any change of units for one of the variables amounts to multiplication of one of these fixed units by a positive number. Suppose we change to a triple of units whose volume measure requires multiplication of liters by $\alpha$, whose temperature measure requires multiplication of kelvin by $\beta$, and whose quantity measure requires multiplication of moles by $\gamma$. Defining the functions $f_1, f_2, f_3 : \mathbb{R}^+ \to \mathbb{R}^+$ by $f_1(x) = \alpha x$, $f_2(x) = \beta x$, and $f_3(x) = \gamma x$, and setting $f = (f_1, f_2, f_3)$, it is appropriate to write the equation relating the variables as

$$P_f(v,t,n) = R(f)\frac{1}{v}nt, \tag{3.2}$$

indicating the particular dependence on the units employed. The functions $f_1$, $f_2$, and $f_3$ are called 'representations,' with each amounting essentially to a choice of unit for a particular variable.[1] Note that, with this notation, Eq. (3.1) would be rewritten $P_{\bar{\iota}}(v,t,n) = R(\bar{\iota})\frac{1}{v}nt$, where $\bar{\iota} = (\iota,\iota,\iota)$ for $\iota$ the identity function (defined by $\iota(x) = x$) on $\mathbb{R}^+$.

A minimal requirement for a law relating physical variables is that the particular choice of representations should not alter the numerical description of the phenomenon in any essential manner. This intuitive notion may be subject to different interpretations; we propose one of them here. Suppose we measure the pressures of an ideal gas at two different triples of volume, temperature, and quantity, using the respective representations $f_1$, $f_2$, and $f_3$, and we find that the first pressure is less than or equal to the second (with the same pressure representation used for each). The relationship between the two pressure computations should hold even if we use different representations $g_1$, $g_2$, and $g_3$. In other words, it should be the case that, for any representations $g_1$, $g_2$, $g_3$, with $g = (g_1, g_2, g_3)$, we have

[1] In this we follow Narens (2002), who uses the term "representation" rather than "scale," which also is in common use (Stevens, 1951).

$$P_f(f(v,t,n)) \leq P_f(f(v',t',n'))$$

(3.3) iff

$$P_g(g(v,t,n)) \leq P_g(g(v',t',n')).$$

Note that the function $P$ in Eq. (3.2) satisfies this requirement. Indeed, with $f_1$, $f_2$, and $f_3$ as above, we have

$$P_f(f(v,t,n)) \leq P_f(f(v',t',n'))$$

iff

$$R(f)\frac{1}{\alpha v}\gamma n\beta t \leq R(f)\frac{1}{\alpha v'}\gamma n'\beta t'$$

iff

$$\frac{1}{v}nt \leq \frac{1}{v'}n't'.$$

As this last equality does not depend on the representations used, Formula (3.3) follows for any functions $f$ and $g$ specifying the representations. We shall say that the function $P$ satisfies the property of 'meaningfulness.' (A precise definition is given as Definition 3.5.)

The intuitively compelling notion of invariance under changes in representation has been described in several ways in the measurement literature, and various approaches have been taken in formulating this notion (see especially

Narens, 2002). One approach has been to describe invariance in terms of functional equations which relate independent variables (and their transformations) to dependent variables (and their transformations); see Luce (1959, 1964); Osborne (1970); Aczél et al. (1986); Kim (1990). In another approach, invariance is described via automorphisms of qualitative structures of nonempty sets and relations on these sets; if certain (additional) constraints are assumed for the structures, strong results which link physical or psychophysical variables may be derived (Luce, 1978, 1990; Falmagne and Narens, 1983; Narens, 2002).

An early formulation, one which may easily have engendered those just mentioned, is due to Suppes and Zinnes:

> A numerical statement is meaningful if and only if its truth (or falsity) is constant under admissible scale transformations of any of its numerical assignments, that is, any of its numerical functions expressing the results of measurement. (Suppes and Zinnes, 1963, p. 66)

(Here, "scale" corresponds to "representation.") This description of meaningfulness is (admittedly) imprecise and may lead not only to more than one approach for its rigorous formulation, but to more than one fundamental interpretation. The equivalence in (3.3) provides one such interpretation: constancy of the truth of a statement is described as a preservation of the order of functional outputs, and admissible transformations are interpreted as being those which match the transformations on which the functions depend. There may be other interpretations of "admissible transformation," however. For instance, consider a fixed $P_f$.

and suppose that there are triples $(v, t, n)$ and $(v', t', n')$ such that

$$P_f(v, t, n) \leq P_f(v', t', n').$$

If for any representations $g_1$, $g_2$, and $g_3$, with $g = (g_1, g_2, g_3)$, we have

$$P_f(g(v, t, n)) \leq P_f(g(v', t', n'))$$

iff

$$P_f(v, t, n) \leq P_f(v', t', n'),$$

then $P_f$ satisfies an invariance property which may be said to satisfy Suppes and Zinnes' description of meaningfulness. However, we shall say in this case that $P_f$ is 'dimensionally invariant.' A formal definition of dimensional invariance is given as Definition 3.6 (see also Causey, 1969; Krantz et al., 1971; Narens, 2002).

Meaningfulness and dimensional invariance are thus seen to be closely related. The two may be hard to separate; indeed, it may seem that any empirical relation that satisfies one must satisfy the other. We will see through the following example that this is not the case.

**Example 3.1.** Choose representations $f_1$ and $f_2$ of length and (positive) temperature difference, respectively, and write $f = (f_1, f_2)$. The final length $L$ of a rod of initial length $\ell$ following an increase $t$ in temperature is given by the equation

$$L_f(\ell, t) = \ell(1 + \zeta(f_2)\, t).$$

in which $\zeta$ is a constant that depends on $f_2$. In particular, if $f_2$ is the representation corresponding to multiplication by $\beta$, then $\zeta(f_2) = \frac{\zeta(\iota)}{\beta}$, where again $\iota$ is the identity function on $\mathbb{R}^+$. Then the function $L_f$ satisfies meaningfulness but not dimensional invariance. (This will be demonstrated below in the Definitions and Basic Concepts section.)

We present a result in Theorem 3.11—the main result of this paper—which ties together the notions of meaningfulness and dimensional invariance. In particular, we show that, under a natural condition relating members of a family of functions, the two notions are equivalent.

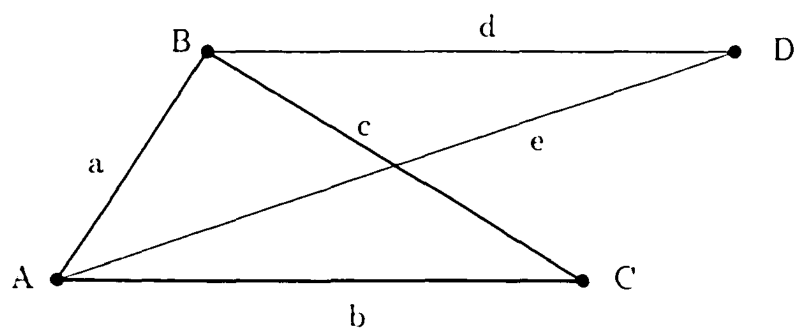


Figure 3.1: Depiction of a transformation that is not factorizable

As mentioned, our main result is a generalization of a result by Falmagne and Narens (1983). The generalization was motivated by close inspection of the types of transformations under which invariance may be studied. Note that each of the transformations considered so far is made up of individual transformations which act independently on separate variables. For instance, the transformation

$f$ considered in Example 3.1 is written $f = (f_1, f_2)$ for the two transformations $f_1$ and $f_2$, each of which acts on a single variable. Such transformations are the ones often considered in the measurement literature (see Narens, 2002). There are important situations, though, in which significant invariances hold under transformations that can not be written as individual transformations on separate variables. For instance, consider the transformation of $\triangle ABC$ to $\triangle ABD$ as shown in Figure 3.1, in which $\overline{BD}$ is constructed parallel to $\overline{AC}$, with the length of $\overline{BD}$ equal to that of $\overline{AC}$. Clearly the area of the triangle is invariant under this transformation. If we define the transformation via the function $f : \prod_{i=1}^{3} ]0,1[ \to \prod_{i=1}^{3} \mathbb{R}^{+}$, where $f(a,b,c) = (a,d,c)$, then there are no functions $f_1, f_2$, and $f_3$ such that $f = (f_1, f_2, f_3)$. In other words, $f$ is not 'factorizable.' (See Definition 3.4 below.)

We give two more examples of transformations which are not factorizable, but under which important invariances hold.

**Example 3.2.** Psychophysicists are interested in the relationships between physical magnitudes of stimuli and the strengths of the sensations they evoke (Fechner, 1860). An important task in psychophysics is the construction of a measure of 'subjective distance' between stimuli based on data which give, for instance, the probability that one stimulus is judged to be different from another. This task, referred to as *Fechnerian scaling*, may be complicated by the fact that the relevant stimuli occupy a multidimensional space. For instance, the stimuli might be auditory tones that vary in both amplitude and frequency. Dzhafarov and Colonius

(2001) propose a theory of Fechnerian scaling which is built in part upon the idea that such distance measures must be invariant with respect to any diffeomorphic transformation of the space of stimulus magnitudes (usually taken to be a subset of $\mathbb{R}^n$). Such transformations may not be factorizable in the multidimensional case.

**Example 3.3.** In the theory of relativity, the "form" of a physical law must be invariant under a particular transformation of the variables called the Lorentz transformation:

> Every general law of nature must be so constituted that it is transformed into a law of exactly the same form when, instead of the space-time variables $x$, $y$, $z$, and $t$ of the original co-ordinate system $K$, we introduce new space-time variables $x'$, $y'$, $z'$, $t'$ of a co-ordinate system $K'$. In this connection the relation between the ordinary and the accented magnitudes is given by the Lorentz transformation. Or in brief: General laws of nature are co-variant with respect to Lorentz transformations. (Einstein, 1961, pp. 42-43).

This transformation is given by

$$(x, y, z, t) \mapsto (x', y', z', t') = \left( \frac{x - \nu t}{\sqrt{1 - (\frac{\nu}{c})^2}}, y, z, \frac{t - \frac{\nu}{c^2} x}{\sqrt{1 - (\frac{\nu}{c})^2}} \right),$$

where $x$, $y$, and $z$ are position coordinates, $t$ is time, $c$ is the speed of light, and $\nu$ is the velocity of coordinate system $K'$ with respect to $K$ (in the direction of the $x$-axis of $K$). It is clear that the transformation is not factorizable.

## 3.1 Definitions and Basic Concepts

Let $X$ be a nonempty set, and let $\mathcal{F} = \{f | f : X \overset{\text{onto}}{\rightarrow} X\}$ be a family of surjective functions mapping $X$ onto itself. For any $f \in \mathcal{F}$, let $M_f$ be a function mapping $X$ to a linearly ordered set $Z$, with the order written $(Z, \leq)$. In the examples above, $X \subseteq \mathbb{R}^n$ and $Z \subseteq \mathbb{R}$.

We call $\mathcal{M} = \{M_f \,|\, f \in \mathcal{F}\}$ a *family of ordinal codes*. Each $M_f \in \mathcal{M}$ is an *ordinal code*.

In this section, we present formally the concepts of meaningfulness and dimensional invariance. We emphasize that the transformations involved may or may not be factorizable. The precise definition of factorizability is as follows:

**Definition 3.4.** Suppose $X = \prod_{i=1}^{n} X_i$ and $Y = \prod_{i=1}^{n} Y_i$ for nonempty sets $X_i$ and $Y_i$, $i = 1, \ldots, n$. A function $f : X \rightarrow Y$ is *factorizable* if there exist functions $f_i : X_i \rightarrow Y_i$, for $i = 1, \ldots, n$, such that $f(x_1, \ldots, x_n) = (f_1(x_1), \ldots, f_n(x_n))$ for all $(x_1, \ldots, x_n) \in X$.

The following two definitions formalize and generalize the concepts of meaningfulness and dimensional invariance introduced earlier through examples.

**Definition 3.5.** The ordinal code $M_f \in \mathcal{M}$ is *meaningful* if, whenever $f^* \in \mathcal{F}$, we have

$$M_f[f(x)] \leq M_f[f(y)]$$

iff

$$M_{f^*}[f^*(x)] \leq M_{f^*}[f^*(y)]$$

for all $x, y \in X$. If this holds for all $M_f \in \mathcal{M}$, we say that $\mathcal{M}$ is *meaningful*.

**Definition 3.6.** The ordinal code $M_f \in \mathcal{M}$ is *dimensionally invariant* if, whenever $f^*, g^* \in \mathcal{F}$, we have

$$M_f[f^*(x)] \leq M_f[f^*(y)]$$

iff

$$M_f[g^*(x)] \leq M_f[g^*(y)]$$

for all $x, y \in X$. If this holds for all $M_f \in \mathcal{M}$, we say that $\mathcal{M}$ is *dimensionally invariant*.

As mentioned, though the notions of meaningfulness and dimensional invariance are related, there exist physical laws which satisfy one but not the other. We return to Example 1, which presents a law that is meaningful but not dimensionally invariant.

**Example 3.1 revisited.** Choose representations $f_1$ and $f_2$ of length and (positive) temperature difference, respectively, and write $f = (f_1, f_2)$. The final length $L$ of a rod of initial length $\ell$ following an increase $t$ in temperature is given by the equation

$$L_f(\ell, t) = \ell(1 + \zeta(f_2)\, t), \tag{3.4}$$

in which $\zeta$ is a constant that depends on $f_2$. In particular, if $f_2$ is the representation corresponding to multiplication by $\beta$, then $\zeta(f_2) = \frac{\zeta(\iota)}{\beta}$.

To see that meaningfulness is satisfied, suppose that the representations $f_1$ and $f_2$ correspond to multiplication by $\alpha$ and $\beta$, respectively. Then

$$L_f(f(\ell, t)) \leq L_f(f(\ell', t'))$$

iff

$$\alpha\ell(1 + \zeta(f_2)\,\beta t) \leq \alpha\ell'(1 + \zeta(f_2)\,\beta t')$$

iff

$$\alpha\ell(1 + \frac{\zeta(\iota)}{\beta}\,\beta t) \leq \alpha\ell'(1 + \frac{\zeta(\iota)}{\beta}\,\beta t')$$

iff

$$\ell(1 + \zeta(\iota)\, t) \leq \ell'(1 + \zeta(\iota)\, t').$$

and this final inequality does not depend on the representations $f_1$ and $f_2$. Thus, $L_f$ is meaningful. Now we show that $L_f$ is not dimensionally invariant. We let $f_1$ correspond to multiplication by 1, $f_2$ correspond to multiplication by $\zeta(\iota)$, $g_1$ correspond to multiplication by 1, and $g_2$ correspond to multiplication by 2.

Setting $\ell = 1$, $\ell' = 2$, $t = 3$, $t' = 1$, and $g = (g_1, g_2)$, we have

$$L_f(\ell, t) = 1(1+3) = 4 \leq 2(1+1) = L_f(\ell', t').$$

but

$$L_f(g(\ell, t)) = 1(1+(2)3) = 7 > 2(1+(2)1) = L_f(g(\ell', t')).$$

This means that $L_f$ is not dimensionally invariant. We note that there actually are several physical laws having the form in Eq. (3.4), including Guy Lussac's Law (for the change in volume of an ideal gas under a temperature change) and the law relating specific heats at constant temperature and volume (see, e.g., Hix and Alley, 1958).

It turns out that the notions of meaningfulness and dimensional invariance are independent: in addition to the function above, which is meaningful but not dimensionally invariant, there exist functions which are dimensionally invariant but not meaningful. As an example, consider the function $M_f : \mathbb{R}^+ \times \mathbb{R}^+ \to \mathbb{R}$ defined by $M_f(x, y) = x + \lambda y$, where $f = (f_1, f_1)$ and $f_1$ corresponds to multiplication by $\lambda$. As shown in Falmagne and Narens (1983) and Roberts (1985), this function is not meaningful, but it is dimensionally invariant. In contrast to Example 3.1, this and other available examples of functions which are not meaningful are hypothetical, i.e., are not necessarily associated to any extant empirical laws. This is not surprising, in view of the compelling argument behind our formulation of meaningfulness. Indeed, this argument probably has been a part of scientists' intuition since long before a definition was formalized.

The following Lemma is a well-known result, so the proof is omitted. (See e.g. Munkres, 1975.)

**Lemma 3.7.** *Suppose $f : X \to Y$, with $X$ and $Y$ ordered sets in the order topology. If $f$ is strictly increasing and surjective, then $f$ is a homeomorphism (i.e., a bicontinuous bijection).*

The next two Propositions are of use in the proof of Theorem 3.11.

**Proposition 3.8.** *The family $\mathcal{M}$ is meaningful if, and only if, for each $f, h \in \mathcal{F}$ there exists a strictly increasing $H_{f,h} : M_h(X) \to M_f(X)$ such that*

$$H_{f,h}(M_h[h(x)]) = M_f[f(x)]$$

*for all $x \in X$. In this case, $H_{f,h}$ also is surjective.*

*Moreover, if $Z$ has the order topology, then $H_{f,h}$ and $H_{f,h}^{-1}$ are continuous.*

PROOF. Choose $f, h \in \mathcal{F}$.

($\Rightarrow$): Suppose $\mathcal{M}$ is meaningful. Define the function $H_{f,h}$ by

$$H_{f,h}(M_h[h(x)]) = M_f[f(x)]$$

for all $x \in X$. Then $H_{f,h}$ is well defined and strictly increasing since $\mathcal{M}$ is meaningful. Also, since $f$ and $h$ map X onto itself, $H_{f,h}$ maps $M_h(X)$ onto $M_f(X)$.

($\Leftarrow$): Let $f^* \in \mathcal{F}$. Suppose $H_{f,h}$ and $H_{f^*,h}$ are as described in the statement of the proposition. Since $H_{f,h}$ and $H_{f^*,h}$ are strictly increasing, we have for all $x, y \in X$,

$$M_f[f(x)] = H_{f,h}(M_h[h(x)]) \leq H_{f,h}(M_h[h(y)]) = M_f[f(y)]$$

iff

$$M_h[h(x)] \leq M_h[h(y)]$$

iff

$$M_{f^*}[f*(x)] = H_{f^*,h}(M_h[h(x)]) \leq H_{f^*,h}(M_h[h(y)] = M_{f^*}[f*(y)].$$

Therefore, we have $M_f[f(x)] \leq M_f[f(y)] \Leftrightarrow M_{f^*}[f^*(x)] \leq M_{f^*}[f^*(y)]$, so $\mathcal{M}$ is meaningful.

The *Moreover* statement is proved with an application of Lemma 3.7. ◻

**Proposition 3.9.** *The family $\mathcal{M}$ is dimensionally invariant if, and only if, for each $M_f \in \mathcal{M}$ and for all $g, g^* \in \mathcal{F}$ there exists a strictly increasing $Q_{f,g,g^*}$ : $M_f(X) \to M_f(X)$ such that*

$$Q_{f,g,g^*}(M_f[g^*(x)]) = M_f[g(x)]$$

*for all $x \in X$. In this case, $Q_{f,g,g^*}$ also is surjective.*

*Moreover, if $Z$ has the order topology, then $Q_{f,g,g^*}$ and $Q^{-1}_{f,g,g^*}$ are continuous.*

We omit the proof, which is similar to that of Proposition 3.8.

The following definition gives a property which provides a link between the two formulations of invariance. This property applies to families of ordinal codes, and it requires that any two codes be related by in a natural way, that is, via a mapping that depends only on the indexing transformations.

**Definition 3.10.** The family of ordinal codes $\mathcal{M}$ is *isotone* if there exists a function $M^* : X \to Z$ such that, for each $M_f \in \mathcal{M}$, we have $M_f = m_f \circ M^*$ for some strictly increasing and surjective $m_f : M^*(X) \to M_f(X)$.

Note that there is no loss of generality in assuming that $M^* = M_h$ for any $h \in \mathcal{F}$. Indeed, if $\mathcal{M}$ is isotone, then $M_h = m_h \circ M^*$ and $M_f = m_f \circ M^*$ for functions $M^*, m_h$, and $m_f$ as in Definition 3.10, with $f \in \mathcal{F}$. But then $M_f = (m_f \circ m_h^{-1}) \circ M_h$, and $m_f \circ m_h^{-1} : M_h(X) \to M_f(X)$ is strictly increasing and surjective.

## 3.2 Main Result

The following theorem, which generalizes Theorem 4 in Falmagne and Narens (1983), specifies the relationship among meaningfulness, dimensional invariance, and isotonicity. In particular, it states that meaningfulness and dimensional invariance are equivalent for isotone families of ordinal codes.

**Theorem 3.11.** *Any two of the properties of meaningfulness, dimensional invariance, and isotonicity imply the third.*

PROOF.

(i) Dimensional invariance and isotonicity imply meaningfulness:

Choose $g^* \in \mathcal{F}$. For any $f \in \mathcal{F}$ and $x \in X$, we have

$$M_f[f(x)] = Q_{f,f,g^*}(M_f[g^*(x)]) \qquad \text{[by Prop. 3.9]}$$

$$= (Q_{f,f,g^*} \circ m_{f,g^*})(M_{g^*}[g^*(x)]) \qquad \text{[by isotonicity]}.$$

Since $Q_{f,f,g^*} \circ m_{f,g^*}$ is strictly increasing, Prop. 3.8 gives that $\mathcal{M}$ is meaningful.

(ii) Meaningfulness and isotonicity imply dimensional invariance:

Suppose $\mathcal{M}$ is meaningful and isotone, and let $M_f, M_h \in \mathcal{M}$. Since $\mathcal{M}$ is meaningful, there exists a strictly increasing $H_{f,h}$ such that

$$M_f[f(x)] = H_{f,h}(M_h[h(x)])$$

for all $x \in X$. Since $\mathcal{M}$ is isotone, there exists a strictly increasing and surjective $m_f$ such that

$$M_f[f(x)] = m_{f,h}(M_h[f(x)])$$

for all $x \in X$. Thus,

$$M_h[f(x)] = m_{f,h}^{-1}(M_f[f(x)]) = (m_{f,h}^{-1} \circ H_{f,h})(M_h[h(x)]), \tag{3.5}$$

where $m_{f,h}^{-1} \circ H_{f,h}$ is strictly increasing.

Let $g \in \mathcal{F}$. We have

$$M_g[f(x)] = m_{g,h}(M_h[f(x)]) \qquad \text{[by isotonicity]}$$

$$= (m_{g,h} \circ m_{f,h}^{-1} \circ H_{f,h})(M_h[h(x)]) \qquad \text{[by Eq. (3.5)]}$$

$$= (m_{g,h} \circ m_{f,h}^{-1} \circ H_{f,h} \circ m_{h,g}^{-1})(M_g[h(x)]) \qquad \text{[by isotonicity]}.$$

where $m_{h,g} \circ m_f^{-1} \circ H_{f,h} \circ m_{h,g}^{-1}$ is strictly increasing. Therefore, by Prop. 3.9, $M_g$ is dimensionally invariant. Since $g \in \mathcal{F}$ is arbitrary, we have that $\mathcal{M}$ is dimensionally invariant.

(iii) Dimensional invariance and meaningfulness imply isotonicity:

Suppose $\mathcal{M}$ is dimensionally invariant and meaningful, and choose $M_h \in \mathcal{M}$. Let $f \in \mathcal{F}$ be arbitrary.

Since $\mathcal{M}$ is meaningful, there exists a strictly increasing and surjective $H_{f,h}$ such that

$$M_f[f(x)] = H_{f,h}(M_h[h(x)])$$

for all $x \in X$.

Since $\mathcal{M}$ is dimensionally invariant, there exists a strictly increasing and surjective $Q_{h,h,f}$ such that

$$M_h[h(x)] = Q_{h,h,f}(M_h[f(x)])$$

for all $x \in X$.

Thus,

$$M_f[f(x)] = (H_{f,h} \circ Q_{h,h,f})(M_h[f(x)])$$

for all $x \in X$, where $H_{f,h} \circ Q_{h,h,f} : M_h(X) \to M_f(X)$ is strictly increasing and surjective. Since $f : X \to X$ is surjective, we have for all $a \in X$ that

$$M_f[a] = (H_{f,h} \circ Q_{h,h,f})(M_h[a]),$$

i.e., $\mathcal{M}$ is isotone. □

## 3.3 Discussion

We have compared the notions of dimensional invariance and meaningfulness in the context of arbitrary transformations on the set of functional inputs. The results in Theorem 3.11 generalize those of Falmagne and Narens (1983), who consider invariance only under transformations which can be factorized and written as strictly increasing, surjective, real-valued functions of real variables. These results state that dimensional invariance and meaningfulness are equivalent for families of functions whose members are related via strictly increasing functions. Such results follow in spirit not only Falmagne and Narens (1983), but also Luce (1978), Narens (2002), who compare similar concepts of invariance.

Putative "laws" which are invariant under the Lorentz transformation are particularly interesting because they may be studied both with respect to this transformation and with respect to changes of representation. It is feasible that some of these may not be invariant under changes in representation, or at least would not satisfy dimensional invariance in the sense of Definition 3.6, when only the changes of representation are considered. Note that a study of such "laws" necessarily involves an approach in which invariance notions (i) are stated with suitable generality for the transformations, and (ii) have families of functions as the objects of interest, rather than single functions, as is the approach typically taken. The formulations in the present paper are appropriate for such a study.

The motivation for this study, and perhaps for any study of properties of invariance, is the investigation of the role of invariance in limiting the possible forms that an empirical law may take. As mentioned, there is a literature which seeks to pinpoint the functional forms which may relate independent and dependent variables that are allowed certain types of representations (e.g. Luce, 1959, 1964; Osborne, 1970; Falmagne and Narens, 1983; Aczél et al., 1986; Kim, 1990). These functions are assumed to satisfy certain invariance properties, and quite often these properties are analogous to the notion of classical dimensional invariance (Luce, 1959, 1964; Osborne, 1970; Aczél et al., 1986). (We specify "classical" because the invariance is assumed for a single function, rather than for a family of functions as in the present paper and in Falmagne and Narens, 1983.) Given the laws presented in Equation (3.4)—established laws which do not satisfy dimensional invariance in the sense of Definition 3.6—it may be necessary to examine further this assumption of invariance in attempting to categorize functions suitable to be empirical laws. We have shown, for instance, that dimensional invariance and meaningfulness are distinct among extant physical laws.

that is, there exist physical laws which satisfy one condition of invariance but not the other. In particular, the law given in Example 3.1 is meaningful but not dimensionally invariant. However, note that this law may naturally be rewritten

$$\Delta L_f(\ell, t) = \ell\,\zeta(f_2)\,t, \tag{3.6}$$

where $\Delta L = L(\ell, t) - \ell$, often the quantity of interest. It is straightforward to show that $\Delta L_f$ in Eq. (3.6) is both meaningful and dimensionally invariant. (In fact, under certain assumptions of differentiability, the transformation $o(\ell) = \ell$ is the only transformation that renders $L(\ell, t) - o(\ell)$ meaningful and dimensionally invariant.) One wonders whether dimensional invariance may be unessential: perhaps a law may always be trivially rewritten in a way that recovers dimensional invariance. This does not appear to be the case, as demonstrated by the following two examples:

**Example 3.12.** The probability $P_f(s, t)$ that an electron will exist at an energy state $s$ at absolute temperature $t$ is given by

$$P_f(s, t) = \frac{1}{1 + e^{\frac{s - \xi(f)}{\kappa(f)\,t}}}, \tag{3.7}$$

where $\xi$ and $\kappa$ are constants which may depend on the representations $f_1$ and $f_2$ of $f = (f_1, f_2)$. (The constant $\kappa$ is Boltzmann's constant, and $\xi$ is the Fermi level energy.) Considerations similar to those used for Example 3.1 may be used to show that $P_f$ is meaningful but not not dimensionally invariant.

**Example 3.13.** The final length $D_f(d, v)$ of a rod of initial length $d$ undergoing a velocity $v$ is given by

$$(3.8) \qquad D_f(d, v) = d\sqrt{1 - \frac{v^2}{c(f)^2}},$$

where $f$ specifies the representations and $c(f)$ is a constant (the speed of light). This physical law, called the Lorentz contraction, also is meaningful but not dimensionally invariant.

It is interesting to note that, though Equations (3.4), (3.7), and (3.8) take diverse forms, the functions $L_f$, $P_f$, and $D_f$ in these equations each may be written in the form

$$(3.9) \qquad V_f(a, b) = F\left[\alpha(f)\, a^{\rho}\, \left(\beta(f)\, b^{\delta} + \gamma(f)\right)\right],$$

in which $\alpha(f)$, $\beta(f)$, $\gamma(f)$, $\rho$, and $\delta$ are constants, $a$ and $b$ are real variables, and $F$ is a strictly monotone function. Examination of these and similar physical laws which are not dimensionally invariant, of whether these laws allow associated formulations which are dimensionally invariant, and of how those associated formulations are obtained are lines of current research. These lines sugest the use of dimensional invariance beyond the typical use in classical physics, i.e., beyond the method of dimensional analysis.

# Chapter 4

# Recasting (the Near-miss to) Weber's Law

In many sensory experiments, the smallest perceptible positive difference $\Delta(x)$ between two stimuli with intensities $x$ and $x+\Delta(x)$ (measured in ratio scale units, e.g., grams for weights, watts/m$^2$ for pure tones) is approximately proportional to $x$. This has been dubbed 'Weber's law' (Fechner, 1860). Some authors (Florentine, 1986; Florentine et al., 1987; see also Narens and Mausfeld, 1992; Narens, 1994) propose that the data described by Weber's law and, more importantly, its subsidiaries—e.g., the so-called 'near-miss to Weber's law'—should be captured directly by $x + \Delta(x)$, which is the actual dependent variable in most experimental situations, rather than through $\Delta(x)$. In fact, a number of researchers have presented data in terms of the measure $x + \Delta(x)$ (e.g. Osman et al., 1980; Scharf and Buus, 1986; Florentine et al., 1987, 1993; Buus and Florentine, 1991; Ozimek and Zwislocki. 1996: Zeng. 1998). Our goal in this paper is to bring new. empirically grounded. theoretical arguments to the debate regarding the two indices $\Delta(x)$ and $x + \Delta(x)$ and the possible models for the corresponding data. Our presentation is organized in the form of three theses.

We begin with an important result involving the power law

$$\Delta(x) = Cx^{\alpha} \tag{4.1}$$

(in which $C$ and $\alpha$ are parameters) used by many researchers to describe systematic deviations from Weber's law. Equation (4.1) often gives a good fit to the data with an estimated exponent $\alpha$ different from 1. However. we will prove that Eq. (4.1) with $\alpha \neq 1$ is inconsistent with another equation which is enforced in those common situations in which data are averaged over order of stimulus presentation in a two-alternative. forced-choice (2AFC) task (or over position in a visual discrimination task). This observation is especially relevant in the field of psychoacoustics. in which Eq. (4.1) is used to fit many pure-tone intensity discrimination data. with estimates of $\alpha$ typically around .9.

Next we show, using empirical results from well-known studies, that many pure-tone intensity discrimination data show a power law growth of $x + \Delta(x)$: the model

$$x + \Delta(x) = Kx^{\beta} \tag{4.2}$$

(in which $K$ and $\beta$ are parameters) also provides a good fit to many data that were originally fit using Eq. (4.1). (The goodness-of-fit values are similar.) The estimated values of $\beta$ for Eq. (4.2), though greater than those obtained for $\alpha$ in Eq. (4.1), are consistently less than 1.

We then derive an important logical consequence of this observation, namely that the value of the parameter $\beta$ must vary systematically with the discrimination criterion in those situations in which the aforementioned averaging of data has taken place. It is easily shown (see our discussion of Thesis 3) that if (4.2) holds and $\beta$ is invariant with the criterion, then necessarily $\beta = 1$, contrary to the results of many studies. Of course, this lack of constancy of $\beta$ may also apply in situations in which no such averaging has taken place (Falmagne et al., 1996).

Although our theoretical results apply in a very general class of psychophysical situations, we present our discussion in the specific context of discriminations of either auditory or visual stimuli varying on a single dimension, with the data collected via a 2AFC task; indeed, it is to such situations that Eq. (4.1) has often been applied (e.g. Guilford, 1932; Hovland, 1938; Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988). For simplicity, we use terminology depicting the comparison of two stimuli separated by a delay—often called a two-interval. forced-choice (2IFC) paradigm in the psychoacoustics literature—but the reader could also keep in mind the comparison of adjacent stimuli (in a visual task). Our notation in the rest of this note is a bit more fastidious than usual. as we indicate the criterion value in the definition of $\Delta(x)$ and also keep track of the order (or position) of stimulus presentation in the 2AFC (cf. Luce and Galanter. 1963: Berliner and Durlach. 1973: Falmagne. 1985). Precise definitions are introduced in the next section.

## 4.1 Definitions and Background

Let $x$ and $y$ denote stimulus intensities[1] measured on a ratio scale. so that $x$ and $y$ are both positive numbers: we do not necessarily assume that $y \geq x$. For convenience. we identify a stimulus and its intensity. The ordered pair $(x, y)$ denotes the presentation of $x$ in the first interval followed by $y$ in the second interval. We write $P(x, y)$ for the probability that $y$ presented in the second interval is judged greater than $x$ in the first interval: then $P(y, x)$ is the probability that $x$ in the second interval is judged greater than $y$ in the first interval. We write

(4.3) $$\xi_\nu(x) = y \quad \text{if and only if} \quad P(x, y) = \nu.$$

[1] 'Intensity' is used as a generic term to indicate physical magnitude of the sensory variable.

In words: $\xi_\nu(x)$ *is the intensity in the second interval judged greater than* $x$ *in the first interval with probability* $\nu$. We call $\xi$ the *sensitivity function* of $P$. In practice, the quantity $\xi_\nu(x)$ can be estimated by standard experimental procedures (e.g. adaptive staircase, stochastic approximation). We restrict consideration to pairs of intensities $(x, y)$ whose discrimination probabilities $P(x, y)$ satisfy $0 < P(x, y) < 1$; in the sequel, we use the phrase 'for all intensities' with these restrictions implied. Also, we assume that $P(x, y)$ is strictly decreasing in its first argument and strictly increasing in its second argument. As is customary, we call a *psychometric function* any function $P(x, \cdot) : y \longmapsto P(x, y)$ assigning, for a fixed intensity $x$, the probability $P(x, y)$ of judging $y$ in the second interval to be greater than $x$ in the first interval. See Figure 4.1 for a summary of the relationships among $P(x, \cdot)$, $\nu$, and $\xi_\nu$. (Ignore $\Delta_\nu(x)$ in the figure for the moment.)

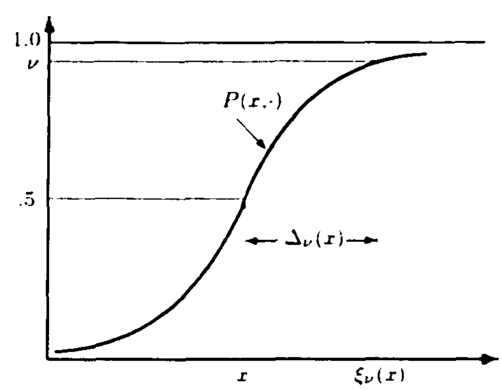


Figure 4.1: A psychometric function $P(x, .)$ and the functions $\Delta_\nu$ and $\xi_\nu$ in a case in which $P(x, x) = .5$.

We turn now to the description of a particular condition on the probabilities $P(x, y)$ which is central to our discussion (see Theses 1 and 3). This condition, which we call the 'balance condition,' may or may not arise naturally in a given psychophysical situation, but, as we argue below, it is automatically enforced as a consequence of the common experimental manipulation of averaging over order in a 2AFC task.

### 4.1.1 The Balance Condition

Our discourse in this subsection will be facilitated by a temporary expansion of our notation. We denote by $P_1(x, y)$ the probability of judging $x$ in the first interval to be greater than $y$ in the second interval, and we denote by $P_2(x, y)$ the probability of judging $y$ in the second interval to be greater than $x$ in the first interval. Note that the function $P_2$ is the same as the function $P$ above. As before, we consider only those pairs $(x, y)$ such that $0 < P_i(x, y) < 1$ (for $i = 1, 2$), using the phrase 'for all pairs $(x, y)$' with these restrictions implied. We assume that $P_1(x, y)$ is strictly increasing in its first argument and strictly decreasing in its second argument, and that $P_2(x, y)$ is strictly decreasing in its first argument and strictly increasing in its second argument. We have, from the definitions of $P_1$ and $P_2$,

$$(4.4) \qquad P_1(x, y) + P_2(x, y) = P_1(y, x) + P_2(y, x) = 1$$

for all pairs $(x, y)$ and $(y, x)$. Note, however, that $P_1(x, y) = P_2(y, x)$ does not necessarily hold in all empirical situations. Indeed, biases based on order or position of stimulus presentation have been observed since Fechner (1860), and order effects in a 2IFC paradigm in psychoacoustics are common (see e.g. Hellström, 1978, 1979). Such effects are often deemed unimportant, or at least are not modeled directly. In particular, experimenters typically allow the stimulus pairs $(x, y)$ and $(y, x)$ to be presented with equal likelihood in a 2IFC task *but do not keep track of listeners' responses separately for the two orderings* (see e.g. Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988) This results in the determination of a single psychometric function which is an average of the two psychometric functions $P_1(\cdot, x)$ and $P_2(x, \cdot)$. In other words, disregarding order information amounts to 'collapsing' the two events

*'$y$ in the second interval is judged greater than $x$ in the first interval'*

*'$y$ in the first interval is judged greater than $x$ in the second interval'*

into a single event

*'$y$ is judged greater than $x$, regardless of order.'*

Essentially, this leads to defining the discrimination probabilities $P(x, y)$ by the equation

(4.5) $$P(x, y) = \frac{P_1(y, x) + P_2(x, y)}{2}.$$

(We direct the reader to Appendix A for two illustrations of Eq. (4.5) arising in practice.)

Together with (4.4), Equation (4.5) leads immediately to

$$P(x,y) + P(y,x) = \frac{P_1(y,x) + P_2(x,y)}{2} + \frac{P_1(x,y) + P_2(y,x)}{2} = 1. \tag{4.6}$$

We refer to the equation

$$P(x,y) + P(y,x) = 1 \tag{4.7}$$

as the *balance condition* (cf. Falmagne, 1985) We emphasize that the balance condition may not hold empirically in a given psychophysical situation: if there are biases based on order of stimulus presentation *and* if no averaging over conditions is performed, then the balance condition will fail. However, whether or not there are biases, if the psychometric function is determined via a method that disregards the order of stimulus presentation, then (4.7) necessarily applies by construction.

### 4.1.2 Weber's Law and the Near-Miss

As can be seen in Figure 4.1, the *Weber function* $\Delta_\nu$ is defined from the sensitivity function $\xi_\nu$ by the equation

$$\Delta_\nu(x) = \xi_\nu(x) - x. \tag{4.8}$$

Weber's law is then expressed by the equation

$$\Delta_\nu(x) = C(\nu)\,x, \tag{4.9}$$

in which the constant of proportionality $C(\nu)$ is strictly increasing with $\nu$. Values adopted for the discrimination criterion $\nu$ typically fall between .70 and .80, with no universal convention (cf. Table 4.1).

As Weber's law is not always satisfied empirically, a number of substitutes have been proposed, a prominent one being the replacement of (4.9) by the power law

$$\Delta_\nu(x) = C(\nu)\, x^{\alpha(\nu)}, \tag{4.10}$$

with $\alpha(\nu) > 0$ a parameter that may depend on the criterion $\nu$. Equation (4.10), which plays a key role in this paper, has been used to fit data from several experimental tasks (see Baird and Noma, 1978), including the judgment of line lengths (Guilford, 1932; Hovland, 1938) and the discrimination of pure tones.

McGill and Goldberg (1968a,b) coined the term 'near-miss to Weber's law' to refer to the fact that, in intensity discriminations between two pure, 1000-Hz tones presented in quiet, Eq. (4.10) seems to hold over a wide range of intensities $x$, with $\alpha$ typically around .9 (see also Riesz, 1928; Dimmick and Olson, 1941). Researchers have examined the effect of a number of experimental conditions on the near-miss, including background noise (e.g. Viemeister, 1972; Moore and Raab, 1974; Hanna et al., 1986; Neff and Jesteadt, 1996), tone frequency (e.g. Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Long and Cullen, 1985; Florentine et al., 1987; Buus and Florentine, 1991; Schroder et al., 1994; Ozimek and Zwislocki, 1996), tone duration (e.g. Green et al., 1979; Florentine, 1986; Buus and Florentine, 1991), tone presentation as continuous or gated (e.g. Green et al., 1979; Viemeister and Bacon, 1988), and hearing ability of the listener (e.g. Florentine et al., 1993; Schroder et al., 1994; Gallégo and Micheyl, 1998). A number of models of loudness coding (see Florentine et al., 1987; Hellman and Hellman, 1990; Allen and Neely, 1997) and some physiological mechanisms (e.g. Gallégo and Micheyl, 1998) have been proposed to account for these effects.

This interest in the near-miss has generated discussion of how best to capture the phenomenon functionally and display it graphically—see Rabinowitz et al. (1976); Jesteadt et al. (1977); Grantham and Yost (1982); Scharf and Buus (1986); Florentine (1986); Florentine et al. (1987); Viemeister and Bacon (1988) and Nelson et al. (1996) in this regard. Especially relevant to our paper is the discussion comparing the two threshold measures

$$10\log(\Delta_\nu(x)/x) \quad \text{and} \quad 10\log(\xi_\nu(x)/x) = 10\log[(x + \Delta_\nu(x))/x].$$

The measure $10\log(\Delta_\nu(x)/x)$ may provide less variability in the threshold estimates at higher thresholds (Jesteadt et al., 1977; Viemeister and Bacon, 1988; Nelson et al., 1996) and is less compressive than $10\log(\xi_\nu(x)/x)$ (Viemeister and Bacon, 1988), while $10\log(\xi_\nu(x)/x)$ is seen as a more direct measure of the intensity discrimination being made (Florentine et al., 1987; see also the Summary and Comments section below) and may be proportional to the sensitivity measure $d'$, allowing the calculation of thresholds corresponding to criteria other than those used empirically (Rabinowitz et al., 1976; Florentine et al., 1987; see also the discussion of Thesis 3 below).

Our aim is not to address directly the issue of which measure—$10\log(\Delta_\nu(x)/x)$ or $10\log(\xi_\nu(x)/x)$—is preferable. Indeed, the choice between the two may depend upon the circumstances and the purpose of the experiment. Rather, we argue that there is a substantive difference between the linear models arising from these measures: one of the models is consistent with the data obtained via averaging, and the other is not. We discuss the inconsistency in our first thesis:

**Thesis 1.** *If both Eq. (4.10) and the balance condition are satisfied (by design or otherwise), then $\alpha(\nu) = 1$ for all $\nu \neq .5$.*

This thesis, which is examined shortly, embodies an obvious admonition against the use of Eq. (4.10) in modeling intensity discrimination data. An alternative model, one that fits many data as well as Eq. (4.10), represents $\xi_\nu$ as a power function:

**Thesis 2.** *Many well-known, pure-tone intensity discrimination data support the hypothesis that $\xi_\nu(x)$ grows as a power law of $x$, that is*

$$\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}, \tag{4.11}$$

*in which $\beta(\nu) > 0$ and $K(\nu) > 0$ are parameters that may depend upon the value $\nu$ of the criterion. In many important cases, the estimated value of the exponent $\beta(\nu)$ in (4.11) is systematically less than 1 (for $.70 < \nu < .80$).*

Thesis 2, useful in that it offers an empirical alternative to Eq. (4.10), also may be used to establish the following:

**Thesis 3.** *For theoretical reasons, the exponent $\beta(\nu)$ in (4.11) must be nonconstant in those situations in which the balance condition holds.*

We consider these theses in turn.

## 4.2 Discussion of Thesis 1

The power law, ubiquitous in psychophysical modeling, has been employed in the form of Eq. (4.10) to describe several empirical situations involving systematic deviations from Weber's law (cf. Baird and Noma, 1978). These situations, which include judgments of line lengths (Guilford, 1932; Hovland, 1938) and discriminations of pure-tone intensities (see the studies in Table 4.1), often give data that are adequately fit by Eq. (4.10), with an exponent $\alpha(\nu)$ less than 1 and greater than about .5. Some authors have questioned the psychological relevance of such a result (e.g. Narens and Mausfeld, 1992), and our arguments, though different from theirs, also question the validity of Eq. (4.10) in these contexts.

As discussed above, frequently the balance condition (4.7) is enforced in the collection or analysis of these data, especially when data are obtained through comparisons of stimulus pairs. It is important to realize that the balance condition greatly limits the scope of empirical situations to which Eq. (4.10) may be applied as a mathematically consistent model. Put another way, Eq. (4.10) obtains as a mathematically consistent model of deviations from Weber's law only when the balance condition does not hold. These facts may be stated precisely as follows:

**Theorem 4.1.** *Suppose that* $\Delta_\nu(x) = C(\nu)\, x^{\alpha(\nu)}$ *holds for all intensities* $x$, *with* $\alpha(\nu) > 0$ *for all* $\nu$ *and* $C(\nu)$ *strictly increasing with* $\nu$. *If* $P(x, y) + P(y, x) = 1$ *for all probabilities* $P(x, y)$, *then* $\alpha$ *is a constant function equal to* 1 *(except possibly at* $\nu = .5$*), i.e.,* $\alpha(\nu) = 1$ *for all criteria* $\nu$ *(except possibly* $\nu = .5$*).*

Thus, the balance condition and Eq. (4.10) are inconsistent with a deviation from Weber's law, insofar as the balance condition and this equation together imply a constant exponent equal to 1.

Notice that the hypotheses involving the positivity of $\alpha$ and the monotonicity of $C$ in Theorem 4.1 are highly plausible empirically. A non-positive value of $\alpha(\nu)$ gives a Weber function measure $\Delta_\nu(x)$ that is not strictly increasing with $x$ for large criteria $\nu$ (say, $\nu > .5$) or not strictly decreasing with $x$ for small $\nu$ (say, $\nu < .5$), which clearly contradict experience in situations to which Eq. (4.10) has been applied.[2] The plausibility of $C(\nu)$ increasing strictly with $\nu$ follows from the fact that $C(\nu)$ equals $\Delta_\nu(1)$ (set $x$ equal to 1 in Eq. (4.10)), which itself should be strictly increasing with $\nu$.

The following is instrumental in our proof of Theorem 4.1:

[2] See however the discussion in the Summary and Comments section regarding the fitting of data which deviate from the near-miss.

**Fact 1.** *Suppose that the equation $Ax^p = x + Bx^q$ holds for all positive real numbers $x$, with $A$ and $B$ nonzero constants. Then $p = q = 1$.*

A proof of Fact 1 appears in Appendix B. Our proof of Theorem 4.1 also relies on the following fact, appearing in Falmagne (1985):

**Fact 2.** *The balance condition holds for all intensities if, and only if, $\xi_\nu$ and $\xi_{1-\nu}$ are inverse functions for all $\nu$.*

To say that $\xi_\nu$ and $\xi_{1-\nu}$ are inverse functions means that

$$\xi_{1-\nu}[\xi_\nu(x)] = x \tag{4.12}$$

for all intensities $x$. (The notation $\xi_{1-\nu}[\xi_\nu(x)]$ means that first the function $\xi_\nu$ is applied to $x$, and then the function $\xi_{1-\nu}$ is applied to the result.) Equation (4.12) arises from the definition of $\xi_\nu$ given in (4.3), which is tantamount to stating that $P(x, \xi_\nu(x)) = \nu$, and from the balance condition, which then gives $P(\xi_\nu(x), x) = 1 - \nu$. This latter expression is equivalent to (4.12) by the definition of $\xi_{1-\nu}$.

We turn now to the proof of Theorem 4.1. Suppose that the balance condition and Eq. (4.10) hold, with $\alpha$ and $C$ having the specified attributes. The balance condition dictates that $P(x, x) = .5$ for all intensities $x$, which gives $\xi_{.5}(x) = x$ by (4.3). Thus, $\Delta_{.5}(x) = 0$ for all $x$, and so $C(.5) = 0$ with $\alpha(.5)$ arbitrary. This explains the parenthetical consideration[3] given $\nu = .5$ in the statement of Theorem 4.1.

[3] As pointed out to us by Geoff Iverson, supposing that $\alpha$ is continuous in $\nu$ (a reasonable assumption) avoids the need for this parenthetical consideration.

We assume for the rest of the proof that $\nu \neq .5$. This implies that $C(\nu)C(1-\nu) < 0$, since $C$ is strictly increasing in $\nu$ and $C(.5) = 0$.

The equation

$$\xi_\nu(x) = x + C(\nu)x^{\alpha(\nu)} \tag{4.13}$$

(arising from (4.8) and (4.10)) and Fact 2 together give the equalities

$$\begin{aligned} x &= \xi_{1-\nu}[\xi_\nu(x)] && \text{[by Fact 2]} \\ &= \xi_\nu(x) + C(1-\nu)[\xi_\nu(x)]^{\alpha(1-\nu)} && \text{[by (4.13)]} \\ &= x + C(\nu)x^{\alpha(\nu)} + C(1-\nu)[x + C(\nu)x^{\alpha(\nu)}]^{\alpha(1-\nu)} && \text{[by (4.13)].} \end{aligned}$$

This implies

$$-C(\nu)x^{\alpha(\nu)} = C(1-\nu)[x + C(\nu)x^{\alpha(\nu)}]^{\alpha(1-\nu)}. \tag{4.14}$$

Since $\alpha(1-\nu) \neq 0$ and $C(1-\nu) \neq 0$, we may rewrite Eq. (4.14) as

$$F(\nu)\, x^{\frac{\alpha(\nu)}{\alpha(1-\nu)}} = x + C(\nu)x^{\alpha(\nu)}, \tag{4.15}$$

where $F(\nu) = \left(\frac{-C(\nu)}{C(1-\nu)}\right)^{\frac{1}{\alpha(1-\nu)}}$. An application of Fact 1 gives $\alpha(\nu) = 1$, and Theorem 4.1 is established.[4]

Theorem 4.1 thus casts doubt on the validity of Eq. (4.10) as a model of deviations from Weber's law. Equation (4.11) provides an alternative model, one without the logical inconsistency of (4.10) (although the balance condition does

[4] As pointed out to us by an anonymous reviewer, from (4.15) and $\alpha(\nu) = 1$ it follows that $C(1-\nu) = \frac{-C(\nu)}{1+C(\nu)}$. This means that under Eq. (4.10) and the balance condition, the psychometric function cannot be symmetric about its point of subjective equality (which would entail $C(1-\nu) = -C(\nu)$).

impose an important and interesting constraint on the exponent in Eq. (4.11) — see the discussion of Thesis 3). Of course, the appropriateness of Eq. (4.11) hinges on its fitting of data which deviate from Weber's law. This topic is examined in the next section.

## 4.3 Discussion of Thesis 2

In many studies, pure-tone intensity discrimination data are plotted as $10\log(\Delta_\nu(x)/x)$ versus $x$ in dB, with the usual observation that the logarithmic transform of Eq. (4.10) provides a good fit over a broad range of intensities and for a variety of experimental conditions (e.g. Jesteadt et al., 1977; Viemeister and Bacon, 1988; Schroder et al., 1994; Neff and Jesteadt, 1996). The data have also been plotted with $\Delta_\nu(x)$ in dB as the ordinate, with of course the same observation; see McGill and Goldberg (1968a,b); Penner et al. (1974); Green et al. (1979); Hanna et al. (1986). As just argued, however, Eq. (4.10) is highly suspect as a model for these data. We propose the simple alternative of Eq. (4.11).

We compare in Table 4.1 the least-squares fits of logarithmic transforms of (4.10) and (4.11) to forty-six data sets from ten well-known studies. In each of the studies, $\Delta_\nu(x)$ was the index used in presenting and analyzing the data, so in particular the parameter and goodness-of-fit estimates in the table were calculated via the equations

$$(4.16) \qquad \log \Delta_\nu(x) = \log(C(\nu)\, x^{\alpha(\nu)}) \qquad \text{[replacing Eq. (4.10)]}$$

and

$$(4.17) \qquad \log \Delta_\nu(x) = \log\left(K(\nu)\, x^{\beta(\nu)} - x\right) \qquad \text{[replacing Eq. (4.11)].}$$

The values presented in the table are in keeping with the original analyses of the data: we averaged over subjects and/or restricted the fits to certain intensities only when done so in the original studies.[5] As indicated in the table, the data from these ten studies cover a wide range of experimental conditions.

Comparison of the root mean square errors indicates that (4.16) and (4.17) fit the data very similarly. Indeed, as illustrated in Figures 4.2 and 4.3, the graph of Eq. (4.11) in dB coordinates for $\Delta_\nu(x)$ and $x$ is very close to that of Eq. (4.10) in these coordinates over the range of intensities tested and for the parameter estimates obtained. Graphs of (4.10) and (4.11) in the coordinates $10\log(\Delta_\nu(x)/x)$ versus $x$ in dB (Figures 4.3 through 4.7) tell a similar story. The nonlinearity of Eq. (4.11) in these coordinates is apparent at very high intensities, however, and though this may be advantageous in fitting some data sets (i.e. Figures 4.6 and 4.7), it is unlikely that (4.11) holds in general at very high intensities. Such is probably the case for (4.10) as well (see especially Viemeister and Bacon, 1988, Figures 1a and 2).

[5] This averaging over subjects is vulnerable to criticism (see the Summary and Comments section below). We employed it only to give an appropriate comparison to the results reported in the original studies.

Table 4.1: Least squares fits of the equations $10 \log \Delta_\nu(x) = 10 \log(C(\nu)\, x^{\alpha(\nu)})$ and $10 \log \Delta_\nu(x) = 10 \log(K(\nu)\, x^{\beta(\nu)} - x)$ to well-known data, along with the corresponding estimates of $\alpha(\nu)$ and $\beta(\nu)$. [a]

| Source | Data Set | Reported $\alpha(\nu)$ | Estimated $\alpha(\nu)$ (RMSE) | Estimated $\beta(\nu)$ (RMSE) |
|---|---|---|---|---|
| McGill & Goldberg[b] (1968a; 1I-2AFC) | Figure 2[c] ($\nu = .75$) | .935 | .919 (1.62) | .984 (1.63) |
| McGill & Goldberg[b] (1968b; 1I-2AFC) | Figure 2[c] ($\nu = .75$) | .905 | .901 (.82) | .973 (.74) |
| Schacknow & Raab | Table 1 | | | |
| (1973; 2IFC) | 250 Hz: S1 | .89 | .89 (.14) | .98 (.36) |
| ($\nu = .75$) | 250 Hz: S2 | .91 | .91 (.71) | .98 (.56) |
| | 1000 Hz: S1 | .87 | .87 (.71) | .97 (.38) |
| | 1000 Hz: S2 | .92 | .92 (.92) | .98 (.79) |
| | 4000 Hz: S1 | .87 | .87 (.94) | .98 (.60) |
| | 4000 Hz: S2 | .87 | .87 (1.27) | .95 (.99) |
| | 7000 Hz: S1 | .81 | .81 (1.32) | .97 (.69) |
| | 7000 Hz: S2 | .88 | .88 (1.46) | .96 (1.21) |
| Penner et al. | Figure 1[d] | | | |
| (1974; 2IFC) | 150 Hz | .86 | .85 (.85) | .92 (.71) |
| ($\nu = .75$) | 250 Hz | .89 | .90 (.94) | .97 (.84) |
| | 1000 Hz | .86 | .87 (.47) | .96 (.43) |
| | 6000 Hz | .88 | .88 (.55) | .95 (.63) |
| | 9000 Hz | .81 | .81 (.36) | .93 (.19) |
| | Figure 2[e] | | | |
| | 9000 Hz: S1 | .82 | .80 (.50) | .96 (.83) |
| | 12000 Hz: S1 | .92 | .90 (.00) | .97 (.04) |
| | 9000 Hz: S2 | .84 | .85 (.24) | .96 (.28) |
| Jesteadt et al. (977; 2IFC) ($\nu = .71$) | Tables B-I and B-II | .928 | .927 (.77) | .987 (.79) |

[a]Both reported and computed $\alpha(\nu)$ values are given to indicate possible inaccuracies in reading graphed data. RMSE stands for root mean square error. 1I-2AFC stands for one-interval, two-alternative forced choice. 2IFC stands for two-interval forced choice, and 3AFC (resp. 3IFC) stands for three-alternative (resp. three-interval) forced choice.

[b]McGill & Goldberg examined just-noticeable decrements in intensity. $\beta(\nu)$ was calculated accordingly.

[c]Restricted to standards greater than 20 dB SL.

[d]Restricted to standards at least 30 dB SL.

[e]Restricted to standards greater than 30 dB SL. Figure 2 also contains another data set, but since it has only two points, it is omitted from this table.

Table 4.1 (continued).

| Source | Data Set | Reported $\alpha(\nu)$ | Estimated $\alpha(\nu)$ (RMSE) | Estimated $\beta(\nu)$ (RMSE) |
|---|---|---|---|---|
| Green et al. | Figure 1[e] | | | |
| (1979; 2IFC) | 10 ms | .867 | .860 (.82) | .948 (.73) |
| ($\nu = .71$) | 100 ms | .884 | .867 (.50) | .984 (.12) |
| | Figure 2[f] | | | |
| | 50 ms | .860 | .850 (.94) | .985 (1.29) |
| | 200 ms | .870 | .875 (.71) | .991 (.97) |
| | 800 ms | .907 | .900 (.47) | .994 (.65) |
| | Figure 3 | | | |
| | 10 ms | .897 | .907 (1.72) | .952 (1.66) |
| | 100 ms | .868 | .887 (.27) | .963 (.48) |
| Hanna et al. | Figure 2[f] | | | |
| (1986; 2IFC) | in quiet | .94 | .94 (2.05) | .98 (2.04) |
| ($\nu = .79$) | in noise | .90 | .89 (1.51) | .97 (1.48) |
| Viemeister & Bacon | Figure 2 | | | |
| (1988; 2IFC) | gated[g] | .92 | .92 (.98) | .98 (.70) |
| ($\nu = .71$) | continuous[h] | .91 | .91 (.57) | .99 (.56) |
| Schroder et al. | Figure 2 (quiet) | | | |
| (1994; 3IFC) | 300 Hz: N1 | .887 | .896 (1.16) | .979 (1.42) |
| ($\nu = .71$) | 300 Hz: N2 | .942 | .953 (1.54) | .991 (1.52) |
| | 300 Hz: N3 | .936 | .938 (1.42) | .989 (1.45) |
| | 500 Hz: N1 | .963 | .960 (1.43) | .992 (1.48) |
| | 500 Hz: N2 | .961 | .947 (1.32) | .992 (1.34) |
| | 500 Hz: N3 | .902 | .908 (1.53) | .982 (1.58) |
| | 1000 Hz: N1 | .934 | .933 (1.54) | .980 (1.60) |
| | 1000 Hz: N2 | .933 | .940 (.68) | .989 (.74) |
| | 1000 Hz: N3 | .889 | .893 (1.41) | .978 (1.69) |
| | 2000 Hz: N1 | .887 | .879 (1.26) | .964 (1.51) |
| | 2000 Hz: N2 | .900 | .907 (.84) | .980 (.91) |
| | 2000 Hz: N3 | .879 | .874 (1.37) | .978 (1.50) |
| | 3000 Hz: N1 | .912 | .890 (1.18) | .970 (1.51) |
| | 3000 Hz: N2 | .906 | .910 (1.26) | .983 (1.46) |
| | 3000 Hz: N3 | .869 | .880 (1.54) | .981 (1.63) |
| Neff & Jesteadt | Figure 2 | .88 | .89 (.41) | .95 (.31) |
| (1996; 3AFC) | | | | |
| ($\nu = .79$) | | | | |

[e]Restricted to standards greater than 30 dB SPL.
[f]Restricted to standards greater than 0 dB SL.
[g]Restricted to standards from 20 to 95 dB SPL.
[h]Restricted to standards from 20 to 85 dB SPL.

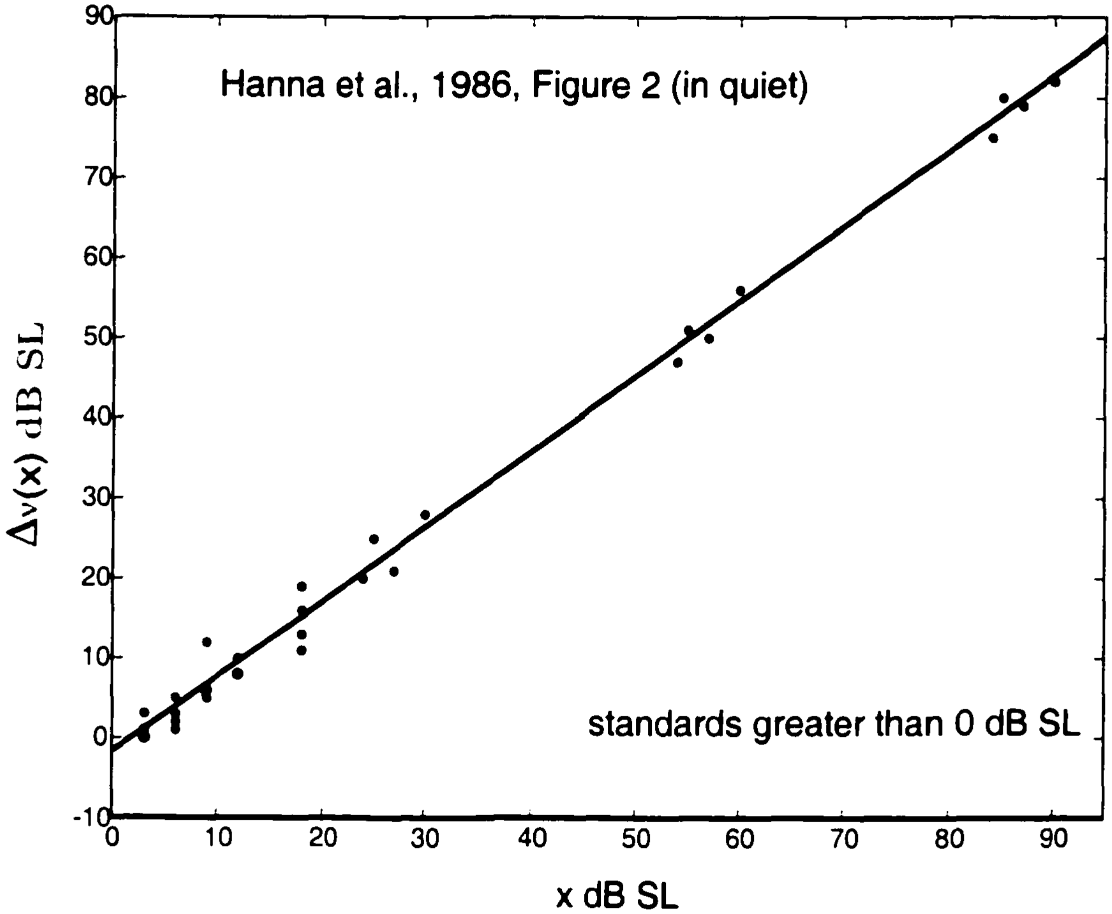


Figure 4.2: Plot of $\Delta_\nu(x) = C(\nu)\,x^{\alpha(\nu)}$ in the coordinates $\Delta_\nu(x)$ dB SL versus $x$ dB SL. The estimates for $C(\nu)$ and $\alpha(\nu)$ were obtained via least squares fit of $10\log\Delta_\nu(x) = 10\log(x^{\alpha(\nu)}C(\nu))$ to the data shown. The root mean square error for this fit is 2.05.

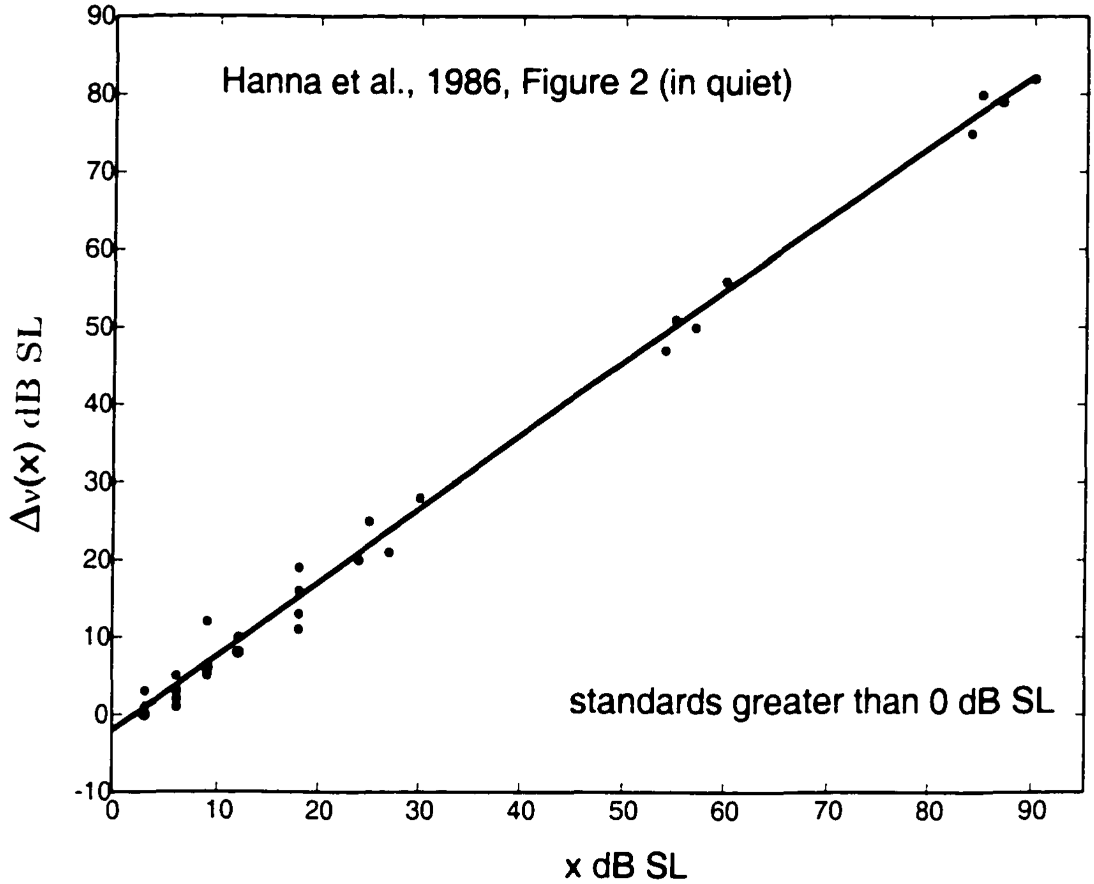


Figure 4.3: Plot of $\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}$ in the coordinates $\Delta_\nu(x)$ dB SL versus $x$ dB SL. The estimates for $K(\nu)$ and $\beta(\nu)$ were obtained via least squares fit of $10 \log \Delta_\nu(x) = 10 \log(x^{\beta(\nu)} K(\nu) - x)$ to the data shown. The root mean square error for this fit is 2.04.

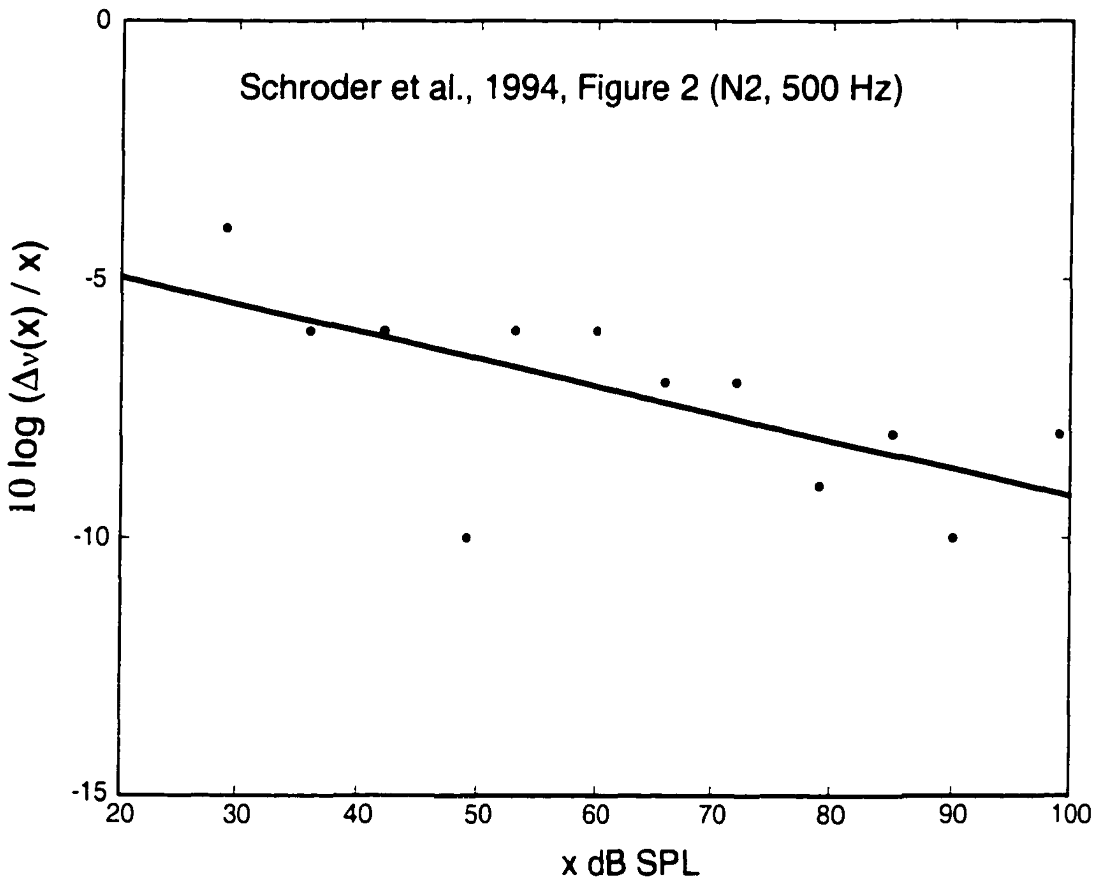


Figure 4.4: Plot of $\Delta_\nu(x) = C(\nu)\,x^{\alpha(\nu)}$ in the coordinates $10\log(\Delta_\nu(x)/x)$ versus $x$ dB SPL. The estimates for $C(\nu)$ and $\alpha(\nu)$ were obtained via least squares fit of $10\log\Delta_\nu(x) = 10\log(x^{\alpha(\nu)}C(\nu))$ to the data shown. The root mean square error for this fit is 1.32.

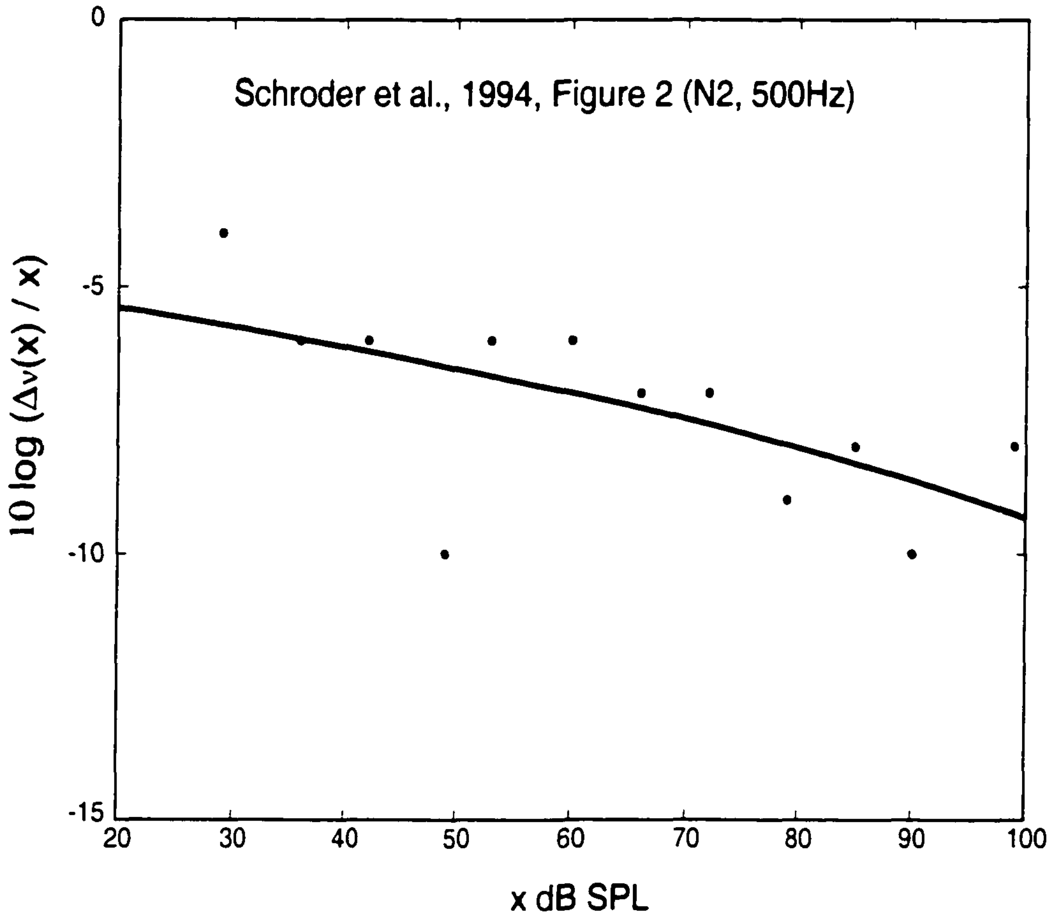


Figure 4.5: Plot of $\xi_\nu(x) = K(\nu)\,x^{\beta(\nu)}$ in the coordinates $10\log(\Delta_\nu(x)/x)$ versus $x$ dB SPL. The estimates for $K(\nu)$ and $\beta(\nu)$ were obtained via least squares fit of $10\log\Delta_\nu(x) = 10\log(x^{\beta(\nu)}K(\nu) - x)$ to the data shown. The root mean square error for this fit is 1.34.

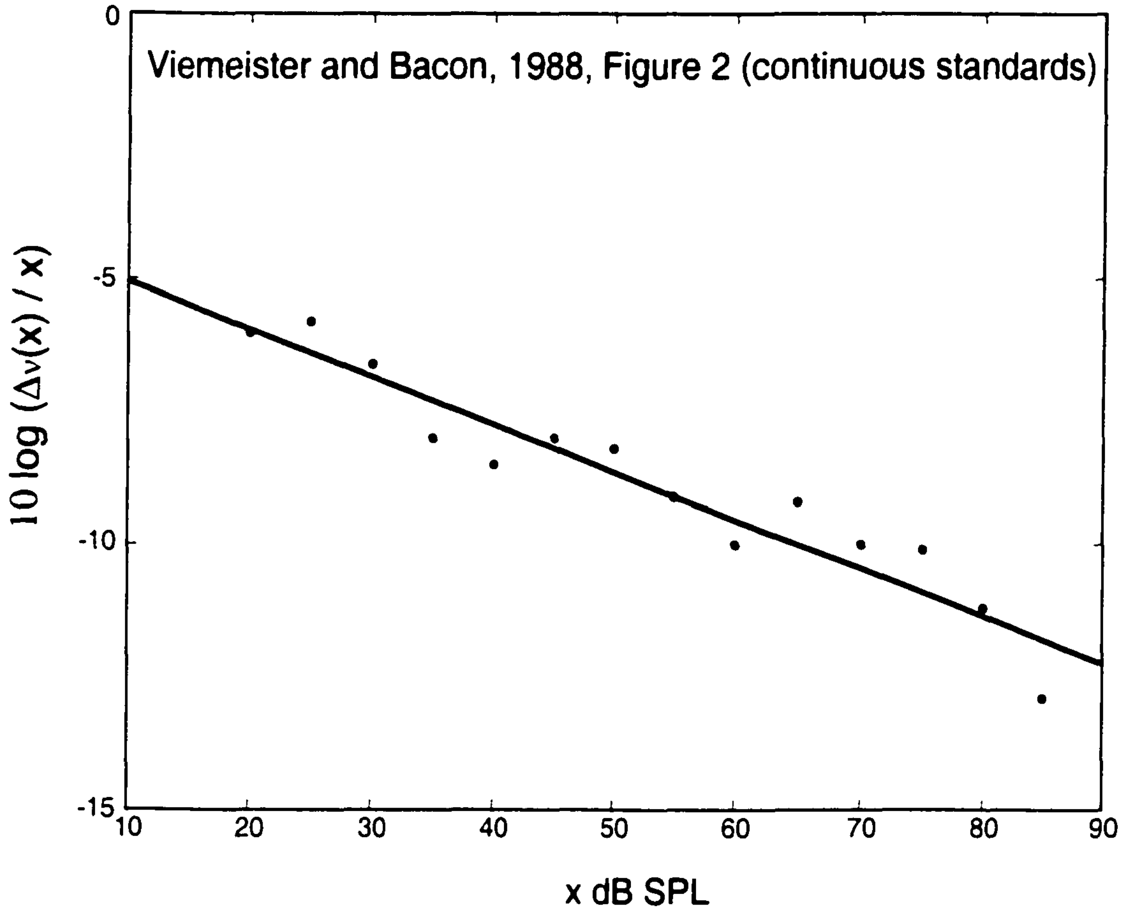


Figure 4.6: Plot of $\Delta_\nu(x) = C(\nu)\, x^{\alpha(\nu)}$ in the coordinates $10\log(\Delta_\nu(x)/x)$ versus $x$ dB SPL. The estimates for $C(\nu)$ and $\alpha(\nu)$ were obtained via least squares fit of $10\log\Delta_\nu(x) = 10\log(x^{\alpha(\nu)}C(\nu))$ to the data shown. The root mean square error for this fit is 0.57.

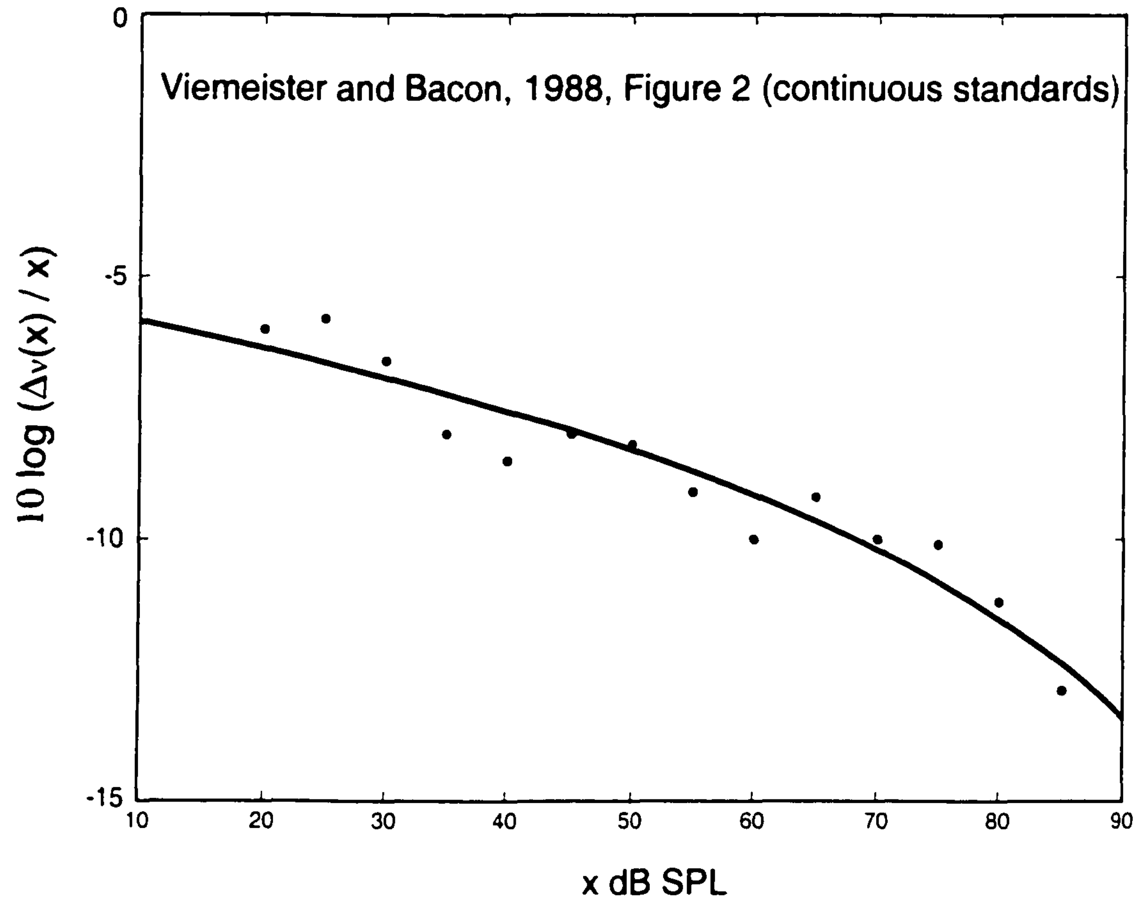


Figure 4.7: Plot of $\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}$ in the coordinates $10\log(\Delta_\nu(x)/x)$ versus $x$ dB SPL. The estimates for $K(\nu)$ and $\beta(\nu)$ were obtained via least squares fit of $10\log \Delta_\nu(x) = 10\log(x^{\beta(\nu)} K(\nu) - x)$ to the data shown. The root mean square error for this fit is 0.56.

The data examined in Table 4.1 support Eq. (4.11) for the wide range of experimental conditions examined. We are not suggesting that this equation describes the data obtained in other experiments, namely those involving extreme stimulus intensities or frequencies, in which the near-miss as modeled by Eq. (4.10) is known to fail (Rabinowitz et al., 1976; Long and Cullen, 1985; Hanna et al., 1986; Florentine et al., 1987; Viemeister and Bacon, 1988; Hellman and Hellman, 1990). Rather, we maintain that the many data adequately described by (4.10) also are adequately described by (4.11), though the latter does not share the logical inconsistency of the former.

We also assert in Thesis 2 that, in many important cases, the estimated value of the exponent in (4.11) is systematically different from 1, and indeed this is true for the data in Table 4.1. However, the estimates of $\beta(\nu)$ in Table 1 are much closer to 1 than the corresponding estimates of $\alpha(\nu)$, and it may be argued that an experimenter could consider these $\beta(\nu)$ estimates as revealing Weber's law without the near-miss. This position would not be justified because the discrepancy is systematic: the estimates of $\beta(\nu)$ are less than 1 in all cases in Table 4.1. Moreover, in view of our argument that $\beta(\nu)$ is nonconstant with $\nu$ (see the discussion of Thesis 3 below), one may very well observe a larger deviation from 1 for different values of $\nu$ (e.g. $\nu$ closer to 1). Data collected in our own laboratory confirm this conjecture (Falmagne et al., 1996; Doble et al., 2002, submitted). It may also be possible to magnify this deviation by increasing the delay between the intervals in which the stimuli appear. (See Hellström, 1979 for a study of this manipulation: see Florentine, 1986 for related analyses.)

The appropriateness of Equation (4.11) for these data is a key component in our discussion of Thesis 3.

## 4.4 Discussion of Thesis 3

Many previous experimenters have, unwittingly, obtained data which establish our Thesis 3, namely, that the exponent $\beta(\nu)$ in (4.11) must be nonconstant with $\nu$. As established earlier, the common manipulation of averaging over order (or position) of stimulus presentation results in the balance condition. This produces a mirror condition on the exponent $\beta(\nu)$ in Eq. (4.11). We will show that—whether or not $\beta(\nu)$ is constant—the balance condition implies

$$\beta(\nu)\beta(1-\nu) = 1 \tag{4.18}$$

for all criterion values $\nu$. This result is well known (cf. Falmagne, 1994). For completeness, we include a proof, which is based on Fact 2 and the following fact:

**Fact 3.** *If Eq. (4.11) holds, then for all criterion values $\nu$ and all intensities $x$ and $\lambda x$ ($\lambda > 0$), we have*

$$\xi_\nu(\lambda x) = \lambda^{\beta(\nu)}\xi_\nu(x). \tag{4.19}$$

Indeed, assuming (4.11), we have successively

$$\xi_\nu(\lambda x) = K(\nu)\,(\lambda x)^{\beta(\nu)} = \lambda^{\beta(\nu)}\,K(\nu)\,x^{\beta(\nu)} = \lambda^{\beta(\nu)}\xi_\nu(x).$$

Equation (4.18) results from the following string of equalities:

$$\begin{aligned}
\lambda x &= \xi_{1-\nu}[\xi_\nu(\lambda x)] && \text{[by Fact 2]} \\
&= \xi_{1-\nu}[\lambda^{\beta(\nu)}\xi_\nu(x)] && \text{[by (4.19)]} \\
&= \lambda^{\beta(\nu)\beta(1-\nu)}\,\xi_{1-\nu}[\xi_\nu(x)] && \text{[by (4.19)]} \\
&= \lambda^{\beta(\nu)\beta(1-\nu)}\,x && \text{[by Fact 2].}
\end{aligned}$$

Dividing by $x > 0$ on both sides yields $\lambda = \lambda^{\beta(\nu)\beta(1-\nu)}$, which in turn gives (4.18) because we can choose $\lambda \neq 1$.

Thus, (4.11) and the balance condition imply (4.18). Note that, in (4.18), if $\beta(\nu) = B$, a positive constant, then $B = 1$. This implies that if $\beta(\nu) \neq 1$ for some value $\nu$ of the criterion, then $\beta$ cannot be a constant function of $\nu$.

We have shown that Thesis 2 holds for data from ten well-known studies. In addition, to the best of our knowledge, the balance condition was enforced in at least six of those studies, viz., those involving 2IFC tasks. Thus, the data from these six studies suggest a nonconstancy of the exponent in (4.11).

It is natural to ask how $\beta(\nu)$ might be expected to vary with $\nu$. The estimated $\beta(\nu)$ values are consistently less than 1 for the studies examined and were obtained for criterion values $\nu$ greater than .5. A glance at Eq. (4.18) reveals that, in these cases, one should expect the estimated values of $\beta(1-\nu)$ to be greater than 1. (Note also that Eq. (4.18) implies $\beta(.5) = 1$.) We have obtained results in our own laboratory, based on extensive data from three subjects, confirming this prediction. In addition, our investigations regarding the specific functional dependence of $\beta$ on $\nu$ suggest that $\beta$ is a decreasing function of $\nu$ (for $.16 < \nu < .84$). One may be able to magnify a deviation from Weber's law simply by increasing the discrimination criterion[6] (Falmagne et al., 1996; Doble et al., 2002, submitted). In any event, we conclude that Thesis 3 is well founded, at least in those situations in which $\beta(\nu)$ is not equal to 1 for at least one value of $\nu$ and the balance condition holds by design or otherwise. This result indicates that, in empirical studies of discrimination, much more attention should be paid to the criterion $\nu$ than has been the case in standard practice.

## 4.5 Summary and Comments

Comparisons of the indices $\Delta_\nu(x)$ and $\xi_\nu(x)$, via the measures $10\log(\Delta_\nu(x)/x)$ and $10\log(\xi_\nu(x)/x)$, are common in the near-miss literature (Jesteadt et al., 1977; Grantham and Yost, 1982; Scharf and Buus, 1986; Florentine et al., 1987; Viemeister and Bacon, 1988), with competing arguments advanced for $\Delta_\nu(x)$ (Jesteadt et al., 1977; Viemeister and Bacon, 1988; Nelson et al., 1996) and for $\xi_\nu(x)$ (Florentine, 1986; Florentine et al., 1987). We have not claimed here that one of the measures is generally preferable to the other, though our results should illuminate the discussion comparing the two. What we have shown is that the near-miss to Weber's law Equation (4.10)—$\Delta_\nu(x) = C(\nu)\,x^{\alpha(\nu)}$ with $\alpha(\nu) \neq 1$—carries a logical inconsistency with a standard empirical technique involving an averaging over

[6] Table 4.1 is not especially helpful in examining this result because of the many experimental conditions represented, including tone frequency, tone presentation as continuous or gated, and intensity range, which likely affect $\beta$.

conditions. We have also presented evidence from a large collection of data that $\xi_\nu(x)$ is, to a good approximation, modeled by the power law Equation (4.11). In the range $.70 \leq \nu \leq .80$, the estimated values of the exponent $\beta(\nu)$ of (4.11), while being greater than those of the exponent $\alpha(\nu)$ of (4.10), remain systematically less than 1. We have derived an important theoretical consequence of this evidence, namely that the exponent $\beta(\nu)$ cannot be constant with the discrimination criterion $\nu$, at least in those situations in which the balance condition is enforced by averaging over order of presentation in a comparison pair. It seems likely that this nonconstancy will occur in other cases as well, and indeed this is confirmed by data collected in our own laboratory (Falmagne et al., 1996; Doble et al., 2002, submitted).

An additional comment concerning $\Delta_\nu(x)$ and $\xi_\nu(x)$ should be made. The relationship between the two is perhaps more subtle than it appears. Note that the graph of Figure 4.1 is predicated on the assumption that $P(x, x) = .5$, or equivalently, that the 'point of subjective equality' of $x$ is equal to $x$, i.e., $\xi_{.5}(x) = x$. Since there are important cases in which this assumption does not hold—for instance, when there are order biases in the 2IFC paradigm in psychoacoustics—it is legitimate to ask whether the Weber function $\Delta_\nu$ should be computed from $\Delta_\nu(x) = \xi_\nu(x) - x$ or from

$$\Delta_\nu(x) = \xi_\nu(x) - \xi_{.5}(x). \tag{4.20}$$

This question is usually answered out of convenience in pure-tone intensity discrimination experiments: the stimuli consist of a 'masker' and a 'masker plus signal.' and $\Delta_\nu(x)$ is determined directly from measurement of the signal. However, we know of no compelling theoretical reason for advocating (4.8) over (4.20) in this situation or in any other in which $\xi_{.5}(x)$ and $x$ may be different.[7] Direct use of the measure $\xi_\nu(x)$ does not require such a choice, which can have an important impact on the interpretation of the data.

As noted, we have compared Equations (4.10) and (4.11) only for near-miss data that were originally described using Eq. (4.10). These data typically were obtained for stimulus levels greater than about 30 dB. However, the near-miss as given by Eq. (4.10) is known to fail at low intensities (e.g. Hanna et al., 1986; Viemeister and Bacon, 1988; Hellman and Hellman, 1990). Substitutes for (4.10) such as

$$(4.21) \qquad \Delta_\nu(x) = C(\nu)[x + x_0(\nu)]^{\alpha(\nu)}$$

have been used to model this fact, in which $x_0(\nu) \geq 0$ is interpreted as an 'internal noise' parameter (Viemeister and Bacon, 1988). Unfortunately, whereas this model gives a good fit to such data, it shares the inconsistency of (4.10): under the balance condition, the exponent $\alpha(\nu)$ in (4.21) must equal 1 for all $\nu \neq .5$. (Arguments nearly identical to those which establish Theorem 4.1 may be used to show this.) A natural generalization of Equation (4.11), the equation

$$(4.22) \qquad \Delta_\nu(x) = K(\nu)[x + x_1(\nu)]^{\beta(\nu)} - x,$$

in which $x_1(\nu) \geq 0$ is an additional parameter, provides a fit to six data sets (from

[7] As pointed out to us by an anonymous reviewer, these considerations may cast doubt on the interpretation of $\Delta_\nu(x)$ as a 'psychological magnitude' (Narens and Mausfeld, 1992).

Viemeister and Bacon, 1988, Figure 3) which is nearly identical to that of (4.21). However, Eq. (4.22) is subject to a similar criticism as Eqs. (4.10) and (4.21): under the balance condition, $x_1(\nu)$ must equal 0 for exponents $\beta(\nu)$ other than 0 or 1. One is left with the task of finding sensitivity functions $\xi_\nu$ which both provide a good fit to these deviations from the near-miss and are consistent with the balance condition, i.e., satisfy (4.12) without ensuing contradiction. Several authors have suggested using a combination of two or more power functions for modeling these and other data which deviate from the near-miss (Rabinowitz et al., 1976; Long and Cullen, 1985; Hanna et al., 1986; Florentine et al., 1987). This should be regarded as a last resort solution. At this point, we leave this matter as an open problem.

Some readers may be puzzled that an inconsistent set of assumptions could yield models that, in the guise of Eqs. (4.10), (4.21), or (4.22), fit many data so well. We wish to make clear that our criticism of these equations is a logical one, not an empirical one. To use a famous historical example, a model describing a planet's orbit in terms of a structure of epicycles may fit astronomical data quite well if the number of epicycles is large, but such a model would be inconsistent with the equations of classical physics and therefore subject to criticism in the Newtonian framework.

Finally, our discussion carries an implicit general warning regarding the averaging over conditions often performed in the controlling for variables regarded as extraneous.[8] Such an averaging is legitimate only in those special circumstances in which the model entertained by the scientist is robust to this averaging, that is, the averaging yields a model of the same form, with different parameters. The same caveat applies of course to the averaging over subjects.

## 4.6 Appendix A

We illustrate with two examples the enforcing of the balance condition vijkna an averaging over conditions. Consider first the experimental situation in which the pairs $(x, y)$ and $(y, x)$ are presented to the participant an equal number of times over the course of many trials, with the participant reporting on each trial which member of the pair appears greater. Identifying probabilities and relative frequencies of responses, we clearly have that (4.4) holds. Now, it could be that there are no order biases for $x$ and $y$, in which case $P_1(x, y) = P_2(y, x)$, and substitution into (4.4) gives the balance condition. However, even if there are order biases, disregarding order information exactly corresponds to determining a single psychometric function $P(x, \cdot)$ that satisfies Eq. (4.6). (If $(x, y)$ and $(y, x)$ are not presented an equal number of times, then the averages of $P_1$ and $P_2$ in Eq. (4.6) are weighted averages, but the balance condition (4.7) is still enforced.)

[8] This tendency is a likely misdeed of the ritualistic teaching of the analysis of variance in the social sciences.

The second illustration involves the common situation in which an up-down method (e.g. Levitt, 1971) is used to determine the psychometric function. A priori, we have the two sensitivity functions $\xi_{1,\nu}$ and $\xi_{2,\nu}$ of $P_1$ and $P_2$, respectively, defined by the two equivalences

$$\xi_{1,\nu}(x) = y \iff P_1(y, x) = \nu \qquad \text{and} \qquad \xi_{2,\nu}(x) = y \iff P_2(x, y) = \nu.$$

In this situation, a probability $\nu$ is fixed along with a value $x$ and, in either of the two orders of presentation, the value $y$ is determined such that $y$ is judged as greater than $x$ with probability $\nu$. In the typical case, $x$ is equally likely to appear in the first or second interval on a given trial, so that $y$ is necessarily between $\xi_{1,\nu}(x)$ and $\xi_{2,\nu}(x)$ (or equal to them, if there is no order bias). As the two psychometric functions $P_1(\cdot, x)$ and $P_2(x, \cdot)$ are linear and parallel in the small region from $\xi_{1,\nu}(x)$ to $\xi_{2,\nu}(x)$, we have that $y$ is simply $(\xi_{1,\nu}(x) + \xi_{2,\nu}(x))/2$, and the point $P(x, y)$ estimated on the single psychometric function is $(P_1(y, x) + P_2(x, y))/2$.

## 4.7 Appendix B

The following is a proof of Fact 1, which states that *if the equation*

$$Ax^p = x + Bx^q \tag{4.23}$$

*holds for all positive real numbers $x$, with $A$ and $B$ nonzero constants, then $p = q = 1$.*

Note that $A = 1 + B$, which is obtained by setting $x = 1$ in (4.23). Choose a positive real number $\lambda$ such that $\lambda \neq 1$, and substitute $\lambda x$ for $x$ in (4.23) to obtain

$$A\lambda^p x^p = \lambda x + B\lambda^q x^q, \tag{4.24}$$

i.e.,

$$\lambda^p(x + Bx^q) = \lambda x + B\lambda^q x^q, \tag{4.25}$$

which upon rearrangement gives

$$(\lambda^p - \lambda)x = Bx^q(\lambda^q - \lambda^p). \tag{4.26}$$

If $p \neq 1$ then we must have $q = 1$, since differentiating both sides of (4.26) with respect to $x$ gives

$$(\lambda^p - \lambda) = Bqx^{q-1}(\lambda^q - \lambda^p), \tag{4.27}$$

and the right-hand side of (4.27) must be nonzero and not varying with $x$ since the left-hand side is nonzero and not varying with $x$. But if $p \neq 1$ and $q = 1$, then from (4.23) we have

$$x^p = \left(\frac{1 + B}{A}\right) x = x, \tag{4.28}$$

and this contradicts the fact that $p \neq 1$. Therefore, we must have $p = 1$, and from (4.26) we must then also have $q = 1$.

# Chapter 5

# Systematic Covariation of the Parameters in the Near-miss to Weber's Law, Pointing to a New Law

## 5.1 Introduction

### 5.1.1 Background

The term 'near-miss to Weber's law,' coined by McGill and Goldberg (1968a,b), refers to the slight but systematic failure of Weber's law occurring for many pure-tone intensity discrimination data. In particular, data obtained for intensity discriminations between pure, 1000-Hz tones presented in quiet show a systematic decrease in the Weber fraction with an increase in intensity, rather than a constant Weber fraction as predicted by Weber's law. More precisely, let us denote by $x$ the intensity of such a tone and by $x + \Delta(x)$ the intensity of a similar tone judged 'just-noticeably' more intense than $x$. The Weber fraction $\frac{\Delta(x)}{x}$ data in these experiments are usually well fitted by one of the two power law models

(5.1) $$\Delta(x) = C\,x^{\alpha} \qquad (C,\ \alpha \text{ are parameters})$$

or

(5.2) $$x + \Delta(x) = K\,x^{\beta} \qquad (K,\ \beta \text{ are parameters}),$$

with the exponents $\alpha$ and $\beta$ estimated to be slightly less than one. Weber's law obtains when $\alpha$ or $\beta$ equal one. The effects of several experimental conditions—including background noise (e.g. Viemeister, 1972; Moore and Raab, 1974; Hanna et al., 1986; Neff and Jesteadt, 1996), tone frequency (e.g. Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Long and Cullen, 1985; Florentine et al., 1987; Buus and Florentine, 1991; Schroder et al., 1994; Ozimek and Zwislocki, 1996), tone duration (e.g. Green et al., 1979; Florentine, 1986; Buus and Florentine, 1991), tone presentation as continuous or gated (e.g. Green et al., 1979; Viemeister and Bacon, 1988), and hearing ability of the listener (e.g. Florentine et al., 1993; Schroder et al., 1994; Gallégo and Micheyl, 1998)—have been examined in the context of these models.[1] These examinations have contributed

[1] Equations (5.1) and (5.2) fit a diversity of pure-tone intensity discrimination data obtained over a rather broad range of stimulus frequencies (say, 250-8000 Hz) and magnitudes (roughly

to the regnant explanation for the near-miss. which is that. while Weber's law may apply for a single auditory 'channel.' an increasing spread of excitation to channels tuned to frequencies above that of the stimulus occurs as stimulus intensity is increased. resulting in improved discriminability at higher intensities (Florentine and Buus. 1981: see also Florentine. 1986. Viemeister and Bacon. 1988. Schroder et al.. 1994). A corpus of recent work incorporates the results of these investigations. along with the spread-of-excitation explanation. in the construction of models of the neural activity driving loudness coding (see Hellman and Hellman. 1990: Allen and Neely. 1997. and the references therein).

Throughout this research. however. little attention has been given the possible effect of subordinate experimental factors. such as the choice of discrimination criterion. on the near-miss. This paper contains considerable data which confirm a theoretical argument that the amount of deviation from Weber's law depends systematically upon the discrimination criterion used. In particular. the data give strong evidence that the exponent $\beta$ in Eq. (5.2) is a (strictly) decreasing function of the criterion. Moreover. the data point toward a systematic covariation of the parameters $\beta$ and $K$ in Eq. (5.2). which suggests a submodel of Eq. (5.2) in which $\beta$ and $K$ are related through a fixed-point property. This submodel is consistent with a notion of top-down control of intensity coding proposed by Parker and Schneider (1994) and Schneider and Parker (1990).

---

30 to 90 dB SL). and via a variety of methods. However. they fail to provide an adequate description of data obtained using extreme stimulus frequencies or intensities (e.g. Rabinowitz et al.. 1976; Long and Cullen. 1985; Hanna et al.. 1986; Florentine et al.. 1987; Viemeister and Bacon. 1988; Hellman and Hellman. 1990).

In what follows, $x$ and $y$ denote tone intensities (measured in ratio scale units, e.g. watts/m$^2$) in a two-alternative, forced choice task. (For convenience in the exposition, a tone and its intensity are identified, so that $x$ and $y$ refer both to tones and to tone intensities.) The tone judged more intense than $x$ with probability exactly equal to $\nu$ is written $\xi_\nu(x)$. The *Weber function* $\Delta$ may be defined in terms of the function $\xi$ by the equation

$$\Delta_\nu(x) = \xi_\nu(x) - x. \tag{5.3}$$

An empirical estimate of the value $\xi_\nu(x)$ will be called an estimated *signal level*. Note that, for sufficiently small $\nu$, this estimate will be less than $x$.

The near-miss often is presented via Eq. (5.1), with estimates of $\alpha$ consistently less than one (e.g. McGill and Goldberg, 1968a,b; Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988; Schroder et al., 1994; Neff and Jesteadt, 1996). However, some authors have questioned the appropriateness of Eq. (5.1) as a model of deviations from Weber's law. For instance, Narens and Mausfeld (1992), using a measurement-theoretic approach, argue against the 'psychological significance' of this equation. Doble et al. (2003) show that, following an averaging over order of stimulus presentation in a two-interval, forced-choice (2IFC) task, the exponent $\alpha$ in Eq. (5.1) is forced mathematically to equal one.[2] This conflicts with the

[2]This averaging typically comprises the estimation of a single signal level for a given referent and a given criterion, with the signal level estimate obtained via a design that allows the referent to appear in the first or second interval with equal probability (e.g. Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988).

typical exponent estimates, which hover around 0.9 (e.g. Schacknow and Raab, 1973; Penner et al., 1974; Jesteadt et al., 1977; Green et al., 1979; Hanna et al., 1986; Viemeister and Bacon, 1988). These results cast doubt on the validity of Eq. (5.1) as a model of deviations from Weber's law, even though this equation provides a good fit for many data. Doble et al. (2003) also show that Eq. (5.2) gives a fit similar to that of Eq. (5.1) for many data originally described using Eq. (5.1), but that Eq. (5.2) does not share the logical inconsistency of Eq. (5.1) arising from its typical parameter estimates. For these reasons, the near-miss is modeled in this paper using Eq. (5.2). (Compare Osman et al., 1980; Scharf and Buus, 1986; Florentine et al., 1987, 1993; Ozimek and Zwislocki, 1996; Zeng, 1998.)

### 5.1.2 Preview

In the following sections, an argument is made on both theoretical and empirical grounds that the parameters $\beta$ and $K$ in the power law model Eq. (5.2) vary systematically with the criterion $\nu$. This model is written as

$$\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}, \tag{5.4}$$

with the implication that $\beta$ and $K$ may be nonconstant functions of $\nu$. It is clear from Eq. (5.4) that $K$ should vary with $\nu$: by definition, $\xi_\nu(1)$ is an increasing function of $\nu$, and $\xi_\nu(1)$ equals $K(\nu)$ in the model. The argument that $\beta$ varies with $\nu$ is outlined shortly. The data strongly support not only the variation of both parameters, but also a specific covariation that suggests a submodel of the form

$$\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}, \tag{5.5}$$

in which $x_*$ and $y_*$ are parameters. Thus, the value $K(\nu)$ in Eq. (5.4) may be written as

$$K(\nu) = y_* x_*^{-\beta(\nu)}. \tag{5.6}$$

The experimental evidence suggests that $x_*$ and $y_*$ are very close in value, with the value corresponding to an intensity near the top of the normal range of hearing.

### 5.1.3 Variation of the Exponent with the Criterion: Theory

Falmagne (1985, 1994) has shown that the near-miss exponent should, in principle, vary with the criterion $\nu$. In particular, he argued that if Eq. (5.4) holds for data averaged over interval order in a 2IFC, then necessarily

$$\beta(\nu)\beta(1-\nu) = 1. \tag{5.7}$$

(See these references for a proof.) An immediate consequence is that, if the exponent $\beta$ is a constant function of $\nu$ under the balance condition, then necessarily $\beta(\nu) = 1$ for all $\nu$. Thus, if it is found that $\beta(\nu) \neq 1$ for at least one value of $\nu$, then $\beta$ must vary with $\nu$.

Doble et al. (2003) examined a large collection of pure-tone, intensity discrimination data from well-known studies which employed an averaging over interval order. They found that Eq. (5.4) provides a good description of these data, i.e., a description similar to that of Eq. (5.1), and that the estimates of $\beta$ are consistently less than one (for criteria near 0.75). Coupled with the result in Eq. (5.7), these analyses imply that $\beta$ is not a constant function of $\nu$. How $\beta$ might be expected to vary with $\nu$ under the balance condition can be seen from Eq. (5.7). Noting that the estimates of $\beta$ are less than one for criteria greater than 0.5, one would expect the estimates of $\beta$ to be greater than one for criteria less than 0.5 (and for Weber's law to hold when $\nu = 0.5$). If $\beta$ is assumed to be monotonic and continuous in $\nu$, then $\beta$ should be a decreasing function of $\nu$.

We know of no previous data from which to compare deviations from Weber's law across criteria. Though different criteria are used by different researchers—values of 0.71, 0.75, and 0.79 all are typical—the variety of experimental conditions employed across studies makes the dependence of the exponent on the criterion difficult to ascertain from the existing literature. The central aim of this study is to examine empirically the possible dependence of $\beta$ on the criterion.

## 5.2 Method

### Experiment 1

Two listeners (identified as AT and S6), with normal audiograms, were paid an hourly rate for participating. Both had extensive experience in psychoacoustic tasks, though neither had previous training in a pure-tone, intensity discrimination task.

Signal levels were estimated for various values of the criterion and for various referent levels, using a 2IFC procedure. On each trial, two 1000-Hz tones were presented in successive intervals. The listener reported which interval contained the louder tone. The trials were divided into blocks of 100, and the listener was given a few minutes of rest between any two blocks. During each sequence of four 100-trial blocks, the referent level (40, 50, 60, 70, or 80 dB SPL) was fixed and the level of the comparison tone was adaptively adjusted using a staircase method (Levitt, 1971). On a trial, the referent level was either in the first or the second interval; we refer to these as *Type 1* and *Type 2* trials, respectively. Twelve independent, adaptive tracks were used for each referent level. Specifically, for each referent level, six values of the criterion—0.16, 0.21, 0.29, 0.71, 0.79, and 0.84—were considered for each of the two trial Types. The schedules for the tracks are given in the appendix.

On each trial, one of these twelve tracks was chosen at random. Each track began with identical intensities in the first and second intervals. Thereafter, the intensity value for a track at the end of a block was the intensity value for the track at the beginning of the next block using the same referent level. The original stepsize of 0.4 dB was decreased to 0.2 dB after the third reversal within a block.

An estimate was computed for each track by averaging the levels at the reversal points in the track's direction, excluding the first ten. Approximately 50-100 reversals (generated by approximately 400 trials) comprise each estimate.[3] The track estimates for each of the two trial Types for a given referent and criterion were averaged to obtain the signal level estimate for that referent and criterion. Overall, there were 30 signal level estimates for each listener, arising from six criteria for each of five referent levels.

A random-block design determined the order of referent level conditions. As mentioned above, four blocks had to be completed for a given referent before the next one was sampled without replacement. Listeners completed approximately 50 blocks for each referent level.

Stimuli were generated digitally, played at a sampling rate of 25 kHz and lowpass filtered at 10 kHz. Sounds were presented diotically over Sennheiser HD-4502 headphones to listeners seated individually in a single-walled IAC sound booth. Programmable attenuators controlled presentation levels. Tones were shaped with 20 ms $cos^2$, onset/offset ramps and presented for 300 ms. Interstimulus intervals (ISIs) were 307 ms. Responses were given via keyboard, and

[3]Due to a programming error, the step size was reset to 0.4 dB at the beginning of each block and changed to 0.2 dB following the third reversal within that block. Negligible effects on threshold estimates are expected because of the large number of reversal levels comprising the average. This programming error was corrected for Experiment 2.

response time was unlimited. A new trial was initiated approximately 500 ms after a response.

## Experiment 2

Two listeners, both with normal audiograms, participated in Experiment 2. One listener, CD (the first author), was experienced in pure-tone, intensity discrimination tasks. The other listener, LN, was naive. The experimental conditions for Experiment 2 were very similar to those for Experiment 1. Only the differences are described here.

In Experiment 2, there were eight adaptive tracks for each referent level, one for each of four criteria and two trial Types. The criteria used were 0.16, 0.29, 0.71, and 0.84. A block was made up of 200 trials, and the referent level was fixed for an entire block.

Three different ISIs were used: 100, 307, and 1000 ms. (The ISI was varied in this experiment as a preliminary study of the effect of ISI on the near-miss parameters.) The ISI was fixed over the course of eight sessions (8000 trials). For Listener CD, the ISIs followed the order 100 ms, 1000 ms, and 307 ms; for Listener LN, they followed the order 1000 ms, 307 ms, and 100 ms.

On each trial, one of the eight tracks was chosen. Tracks involving criteria of 0.16 or 0.84 were twice as likely to be chosen as those involving criteria of 0.29 or 0.71.[4] Each track began at a level predicted from Experiment 1. The stepsize was 0.2 dB throughout.

There were 60 different signal level estimates for each listener: one estimate for each of four criteria, five referent levels, and three ISIs. Estimates were obtained after excluding the first four reversals in a track's direction (and then averaging the estimates from the two track Types, as in Experiment 1). Roughly 30-50 reversals, generated by 100-300 trials, comprise each estimate.

The stimuli were generated as in Experiment 1. The intertrial interval was 750 ms for all trials.

## 5.3 Results

The data to be analyzed are the estimated values of the function $\xi_\nu$ for each of the four listeners. For Listeners AT, S6, and CD, the estimated 'level differences' $10\log[\frac{\xi_\nu(x)}{x}]$, for $\nu = 0.71$, fall between about 0.30 and 2.01. These values are in good agreement with those obtained in previous experiments using this criterion (e.g., Jesteadt et al., 1977; Ozimek and Zwislocki, 1996; Florentine, 1986; Florentine et al., 1987). The estimated level differences for $\nu = 0.79$ also are consistent with those of previous studies using this criterion, and range from about 1.0 to

[4]This was done to improve the efficiency of the data collection procedure, as it allowed the number of reversals for the tracks with more extreme criteria to be roughly equal to the number of reversals for the middle criteria.

2.7 (e.g., Hanna et al., 1986; Neff and Jesteadt, 1996). For listener LN, the level differences are slightly higher (0.76 to 3.97 for $\nu = 0.71$), though not without precedent (see e.g. Florentine et al., 1987, listener RS). The experimental methods used in the present study differ slightly from those of many previous studies, in that in the present study, listeners were not given feedback, and the signal level estimates were obtained by averaging track values from the two trial Types, rather than by computing a single value from a track that includes both trial Types. No feedback was given because it is unclear that feedback would serve a purpose in this experimental design: the intensity of the comparison tone may be greater or less than that of the referent on a given trial. Data were recorded from both trial Types to investigate possible differences in the parameter values for the two situations 'referent in the first interval' and 'referent in the second interval.' These investigations will be detailed in future work. In view of the good agreement with previous studies for the data of three of the listeners, it is unlikely that the lack of feedback or the particular choice of data averaging had an effect on the results reported in this paper. We cannot explain the greater variability of the data of Listener LN compared to the data of the other listeners in this study.

The logarithmic transform of Eq. (5.4) was fit to the data for each value of $\nu$ by regressing $10 \log \xi_\nu(x)$ on $10 \log x$. A summary of the fits and parameter estimates appears in columns 3 through 5 of Table 5.1. The prominent feature of these values is the clear, systematic variation of $\beta(\nu)$ with $\nu$: there is only one exception to the strict decrease of $\beta(\nu)$ with $\nu$. Furthermore, the estimates follow the pattern $\beta(\nu) < 1$ for $\nu > 0.5$ and $\beta(\nu) > 1$ for $\nu < 0.5$. These observations strongly support the variation of $\beta$ predicted from Eq. (5.7). It also is clear from the table that estimates of $K(\nu)$ strictly increase as $\nu$ increases, as is suggested by Eq. (5.4).

The fit of the linear model to the data generally is very good. The root mean square errors are reported in column 5 of Table 5.1. Figure 1 displays the data sets and the best-fitting lines for each. Though there are some possible deviations from linearity, these deviations are small and not systematic. The data for Listener LX are more variable than those of the other listeners, but her data are appropriately fit with a linear model, as seen in the figure.

Table 5.1: Estimated values of the parameters in the models $\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}$ and $\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}$. The estimates were obtained using the logarithmic transforms of the respective equations. RM stands for root mean square error, and $\nu$ represents the discrimination criterion. The table is continued on the next page.

| | | $\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}$ | | | $\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}$ | | | |
|---|---|---|---|---|---|---|---|---|
| $\nu$ | | $\beta(\nu)$ | $10 \log K(\nu)$ | RM | $\beta(\nu)$ | $10 \log x_*$ | $10 \log y_*$ | RM |
| 0.16 | AT | 1.028 | -3.329 | 0.141 | 1.028 | 117.3 | 117.3 | 0.141 |
| 0.21 | | 1.026 | -3.165 | 0.163 | 1.027 | | | 0.164 |
| 0.29 | | 1.018 | -2.190 | 0.085 | 1.018 | | | 0.085 |
| 0.71 | | 0.987 | 1.621 | 0.146 | 0.985 | | | 0.152 |
| 0.79 | | 0.982 | 2.363 | 0.095 | 0.977 | | | 0.134 |
| 0.84 | | 0.965 | 3.671 | 0.162 | 0.971 | | | 0.201 |
| 0.16 | S6 | 1.073 | -7.383 | 0.547 | 1.071 | 108.8 | 109.2 | 0.549 |
| 0.21 | | 1.066 | -6.698 | 0.231 | 1.065 | | | 0.232 |
| 0.29 | | 1.040 | -4.113 | 0.198 | 1.043 | | | 0.207 |
| 0.71 | | 0.973 | 2.973 | 0.157 | 0.981 | | | 0.222 |
| 0.79 | | 0.970 | 3.799 | 0.305 | 0.968 | | | 0.307 |
| 0.84 | | 0.965 | 4.518 | 0.306 | 0.960 | | | 0.325 |
| 0.16 | CD | 1.020 | -2.090 | 0.120 | 1.020 | 111.0 | 111.1 | 0.120 |
| 0.29 | 0.1 s | 1.012 | -1.202 | 0.066 | 1.013 | | | 0.067 |
| 0.71 | | 0.995 | 0.756 | 0.069 | 0.994 | | | 0.075 |
| 0.84 | | 0.984 | 1.831 | 0.087 | 0.985 | | | 0.090 |
| 0.16 | CD | 1.013 | -1.693 | 0.135 | 1.014 | 124.3 | 124.2 | 0.135 |
| 0.29 | 0.3 s | 1.006 | -0.762 | 0.074 | 1.006 | | | 0.074 |
| 0.71 | | 0.993 | 0.954 | 0.157 | 0.992 | | | 0.157 |
| 0.84 | | 0.986 | 1.724 | 0.056 | 0.986 | | | 0.057 |
| 0.16 | CD | 1.018 | -2.174 | 0.148 | 1.021 | 115.1 | 115.1 | 0.155 |
| 0.29 | 1 s | 1.015 | -1.542 | 0.037 | 1.013 | | | 0.061 |
| 0.71 | | 0.992 | 1.060 | 0.150 | 0.991 | | | 0.152 |
| 0.84 | | 0.980 | 2.279 | 0.040 | 0.981 | | | 0.047 |

Table 5.1 (continued).

| | | $\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}$ | | | $\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}$ | | | |
|---|---|---|---|---|---|---|---|---|
| $\nu$ | | $\beta(\nu)$ | $10 \log K(\nu)$ | RM | $\beta(\nu)$ | $10 \log x_*$ | $10 \log y_*$ | RM |
| 0.16 | LN | 1.133 | -13.999 | 1.138 | 1.126 | 105.9 | 105.6 | 1.146 |
| 0.29 | 0.1 s | 1.053 | -6.157 | 0.977 | 1.060 | | | 0.986 |
| 0.71 | | 0.946 | 5.128 | 0.212 | 0.954 | | | 0.259 |
| 0.84 | | 0.896 | 11.127 | 0.520 | 0.889 | | | 0.541 |
| 0.16 | LN | 1.109 | -11.417 | 0.485 | 1.110 | 109.7 | 110.3 | 0.485 |
| 0.29 | 0.3 s | 1.070 | -6.966 | 0.615 | 1.067 | | | 0.617 |
| 0.71 | | 0.959 | 4.761 | 0.263 | 0.965 | | | 0.281 |
| 0.84 | | 0.912 | 10.410 | 0.344 | 0.909 | | | 0.349 |
| 0.16 | LN | 1.079 | -10.174 | 1.268 | 1.099 | 127.8 | 129.1 | 1.323 |
| 0.29 | 1 s | 1.079 | -6.676 | 0.485 | 1.049 | | | 0.739 |
| 0.71 | | 0.971 | 4.824 | 0.518 | 0.973 | | | 0.520 |
| 0.84 | | 0.928 | 9.963 | 0.577 | 0.935 | | | 0.593 |

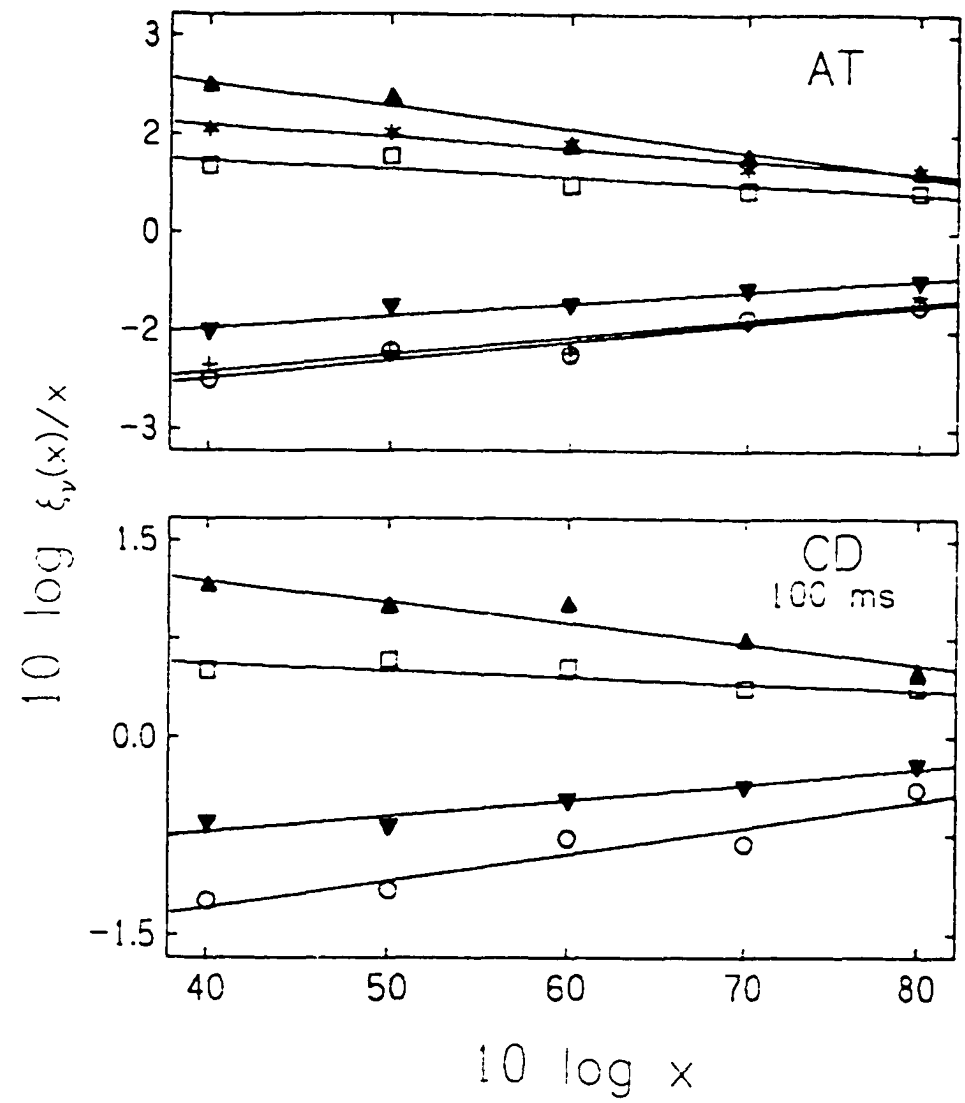


Figure 5.1: Plots of intensity discrimination data for listeners AT and CD (100 ms). The symbols represent different criteria $\nu$: upright triangles for $\nu = 0.16$, stars for $\nu = 0.21$, squares for $\nu = 0.29$, inverted triangles for $\nu = 0.71$, crosses for $\nu = 0.79$, and circles for $\nu = 0.84$. The best-fitting lines were obtained by regression of $10 \log \xi_\nu(x)$ on $10 \log x$.

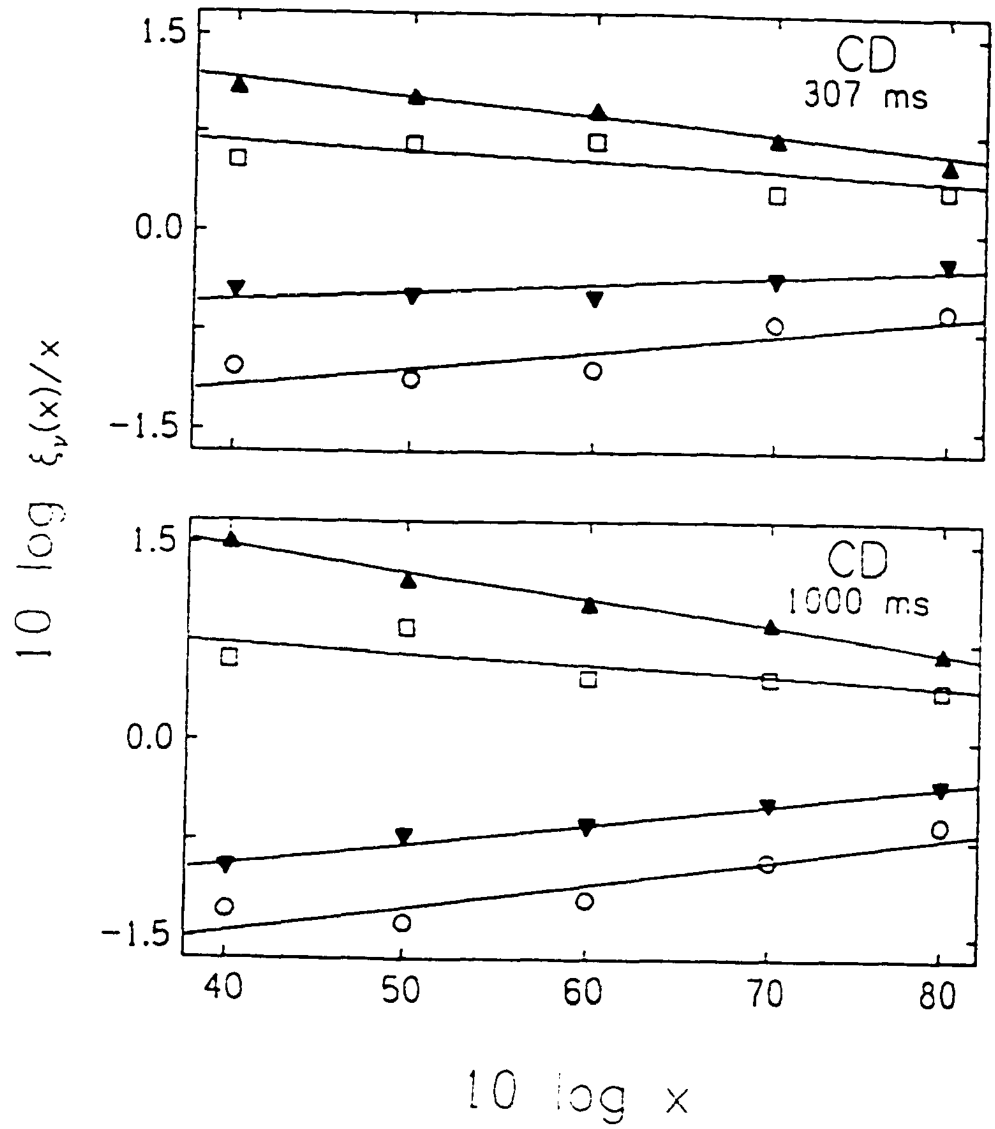


Figure 5.2: Plots of intensity discrimination data for listener CD (307 ms and 1000 ms). The symbols represent different criteria $\nu$: upright triangles for $\nu = 0.16$, stars for $\nu = 0.21$, squares for $\nu = 0.29$, inverted triangles for $\nu = 0.71$, crosses for $\nu = 0.79$, and circles for $\nu = 0.84$. The best-fitting lines were obtained by regression of $10 \log \xi_\nu(x)$ on $10 \log x$.

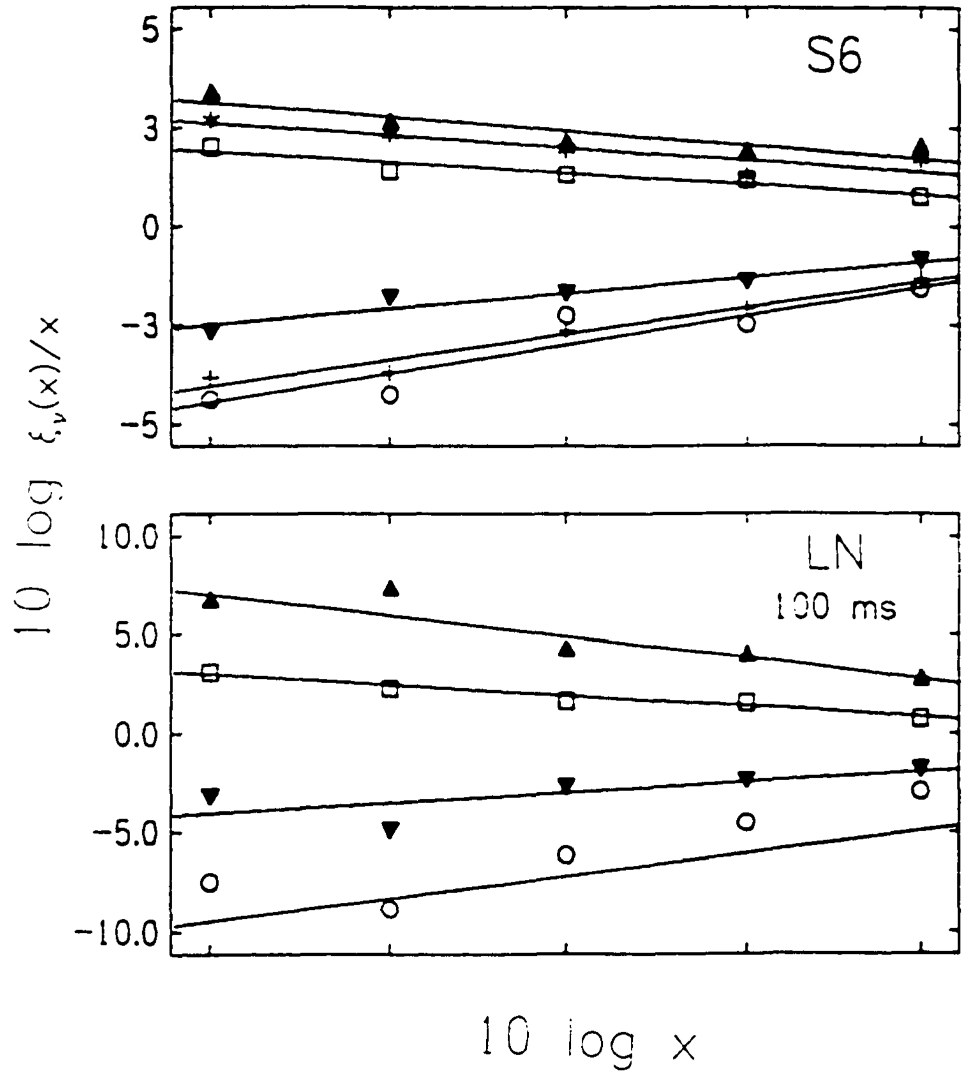


Figure 5.3: Plots of intensity discrimination data for listeners S6 and LN (100 ms). The symbols represent different criteria $\nu$: upright triangles for $\nu = 0.16$, stars for $\nu = 0.21$, squares for $\nu = 0.29$, inverted triangles for $\nu = 0.71$, crosses for $\nu = 0.79$, and circles for $\nu = 0.84$. The best-fitting lines were obtained by regression of $10 \log \xi_\nu(x)$ on $10 \log x$.

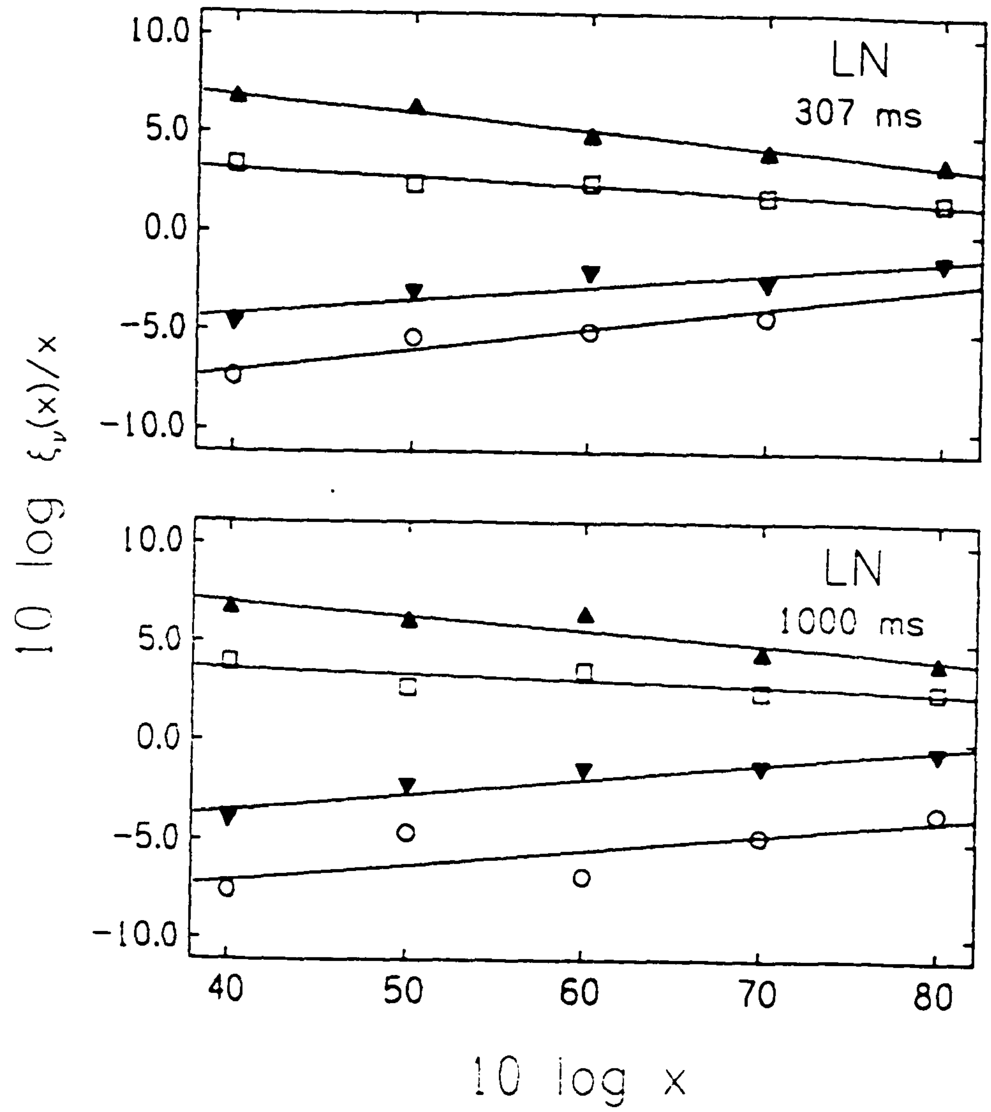


Figure 5.4: Plots of intensity discrimination data for listener LN (307 ms and 1000 ms). The symbols represent different criteria $\nu$: upright triangles for $\nu = 0.16$, stars for $\nu = 0.21$, squares for $\nu = 0.29$, inverted triangles for $\nu = 0.71$, crosses for $\nu = 0.79$, and circles for $\nu = 0.84$. The best-fitting lines were obtained by regression of $10 \log \xi_\nu(x)$ on $10 \log x$.

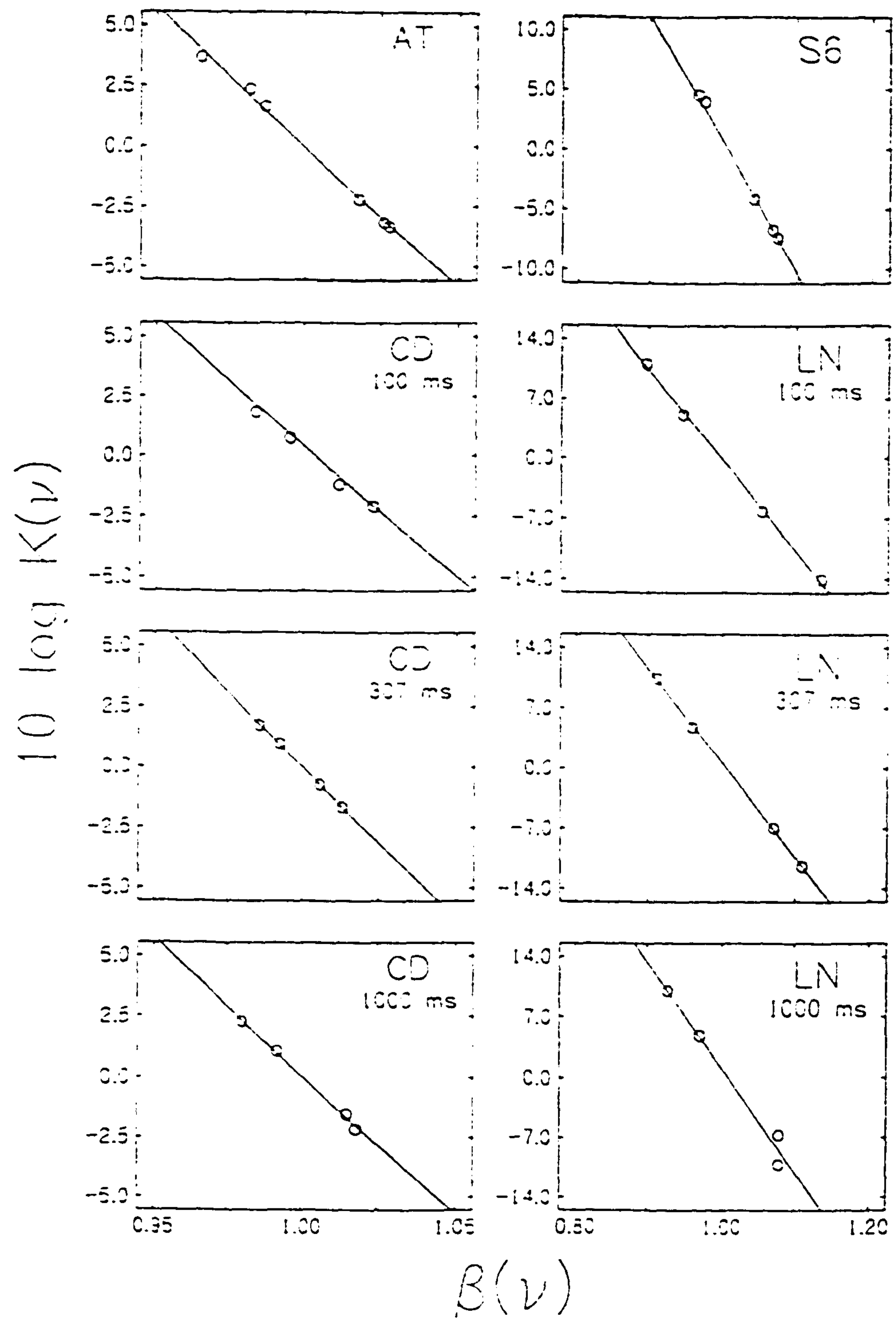


Figure 5.5: Plots of $10 \log K(\nu)$ versus $\beta(\nu)$, with $K(\nu)$ and $\beta(\nu)$ estimated from the model $\xi_\nu(x) = K(\nu)\, x^{\beta(\nu)}$.

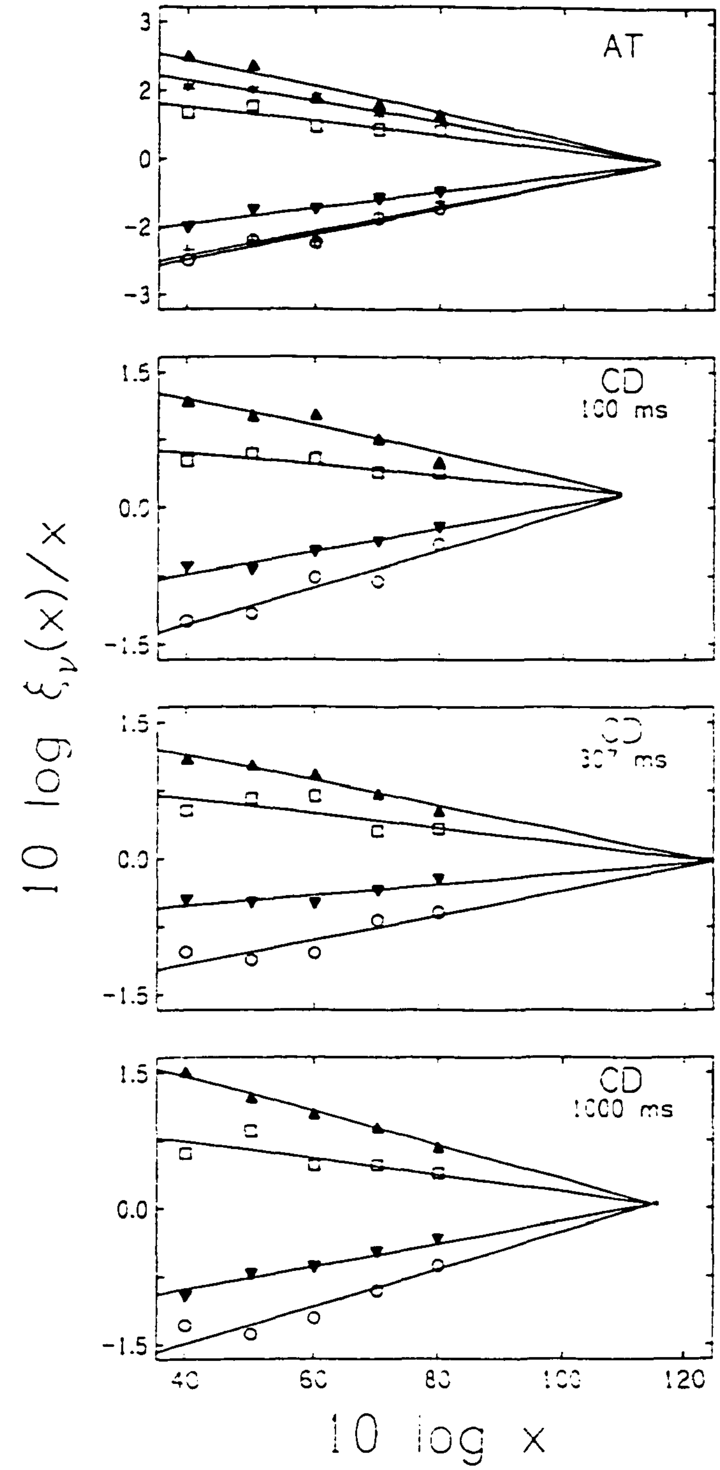


Figure 5.6: Plots of best-fitting lines obtained using the model $\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}$. The data points are the same as those in Figures 5.1 and 5.2, and the symbols have the same meaning as in those figures.

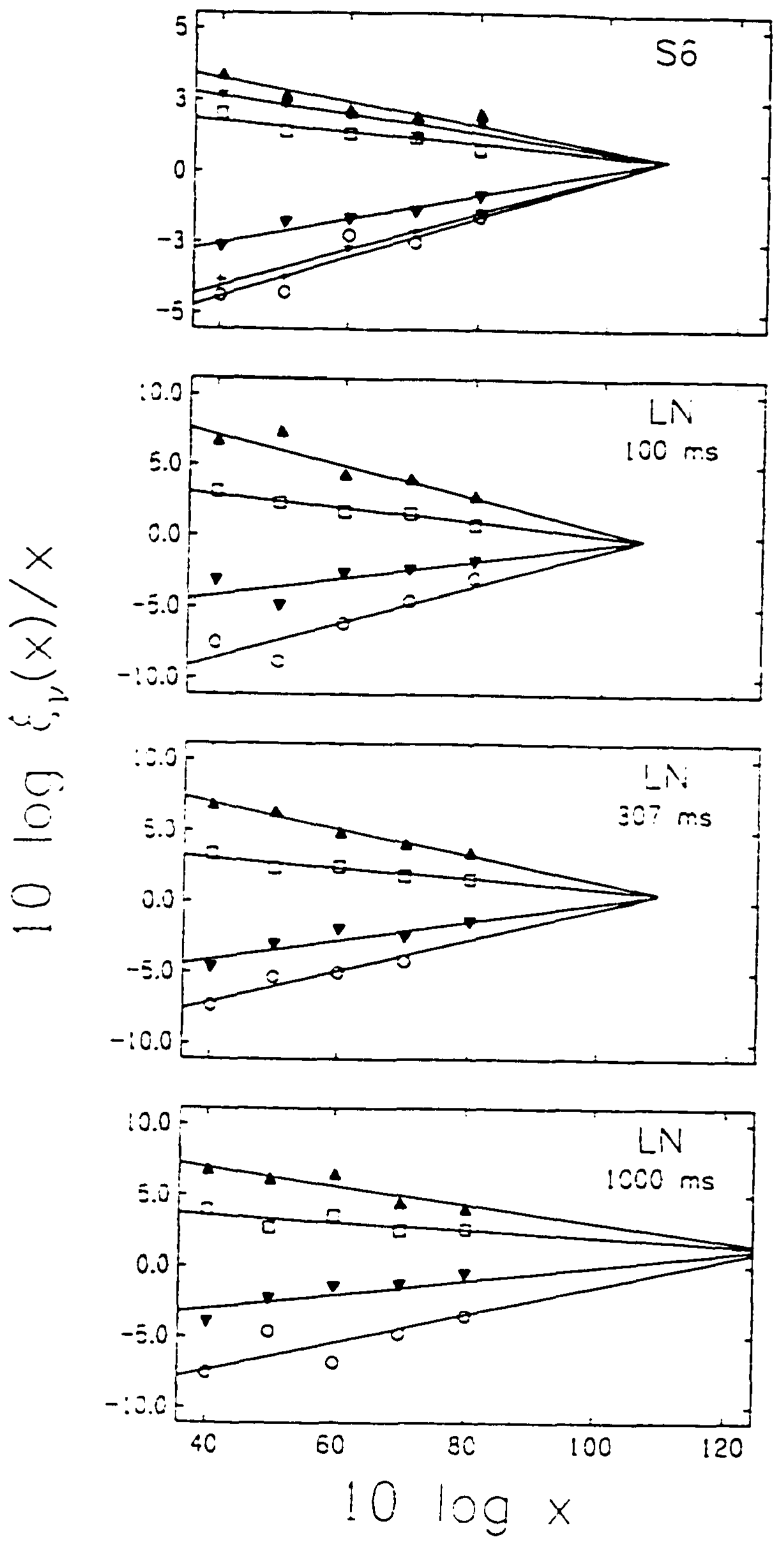


Figure 5.7: Plots of best-fitting lines obtained using the model $\frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}$. The data points are the same as those in Figures 5.3 and 5.4, and the symbols have the same meaning as in those figures.

Studying Figures 5.1-5.4 convinces that much more is to be culled from the data than simply the nonconstancy of the parameters $\beta(\nu)$ and $K(\nu)$. In each of the eight plots in these figures, the best-fitting lines seem to intersect at a common point, suggesting a strong covariation between the parameters $K(\nu)$ and $\beta(\nu)$. The covariation is seen directly in Figure 5.5: plots of $10\log K(\nu)$ versus $\beta(\nu)$ are linear. These observations suggest that $K(\nu)$ and $\beta(\nu)$ are related through the equation

$$(5.8) \qquad 10\log K(\nu) = (-10\log x_*)\,\beta(\nu) + 10\log y_*,$$

with $-10\log x_*$ the slope and $10\log y_*$ the $y$-intercept in Figure 5.5. Solving Eq. (5.8) for $K(\nu)$ gives $K(\nu) = x_*^{-\beta(\nu)}\,y_*$, so that the near-miss model (5.4) may be specialized into

$$(5.5) \qquad \frac{\xi_\nu(x)}{y_*} = \left(\frac{x}{x_*}\right)^{\beta(\nu)}.$$

Note that, for all values of $\nu$ in Eq. (5.5), the graphs of $10\log[\frac{\xi_\nu(x)}{x}]$ versus $10\log x$ pass through the same point $(10\log x_*, 10\log y_* - 10\log x_*)$. This is the property suggested by the convergence of the lines in Figures 5.1-5.4. These lines were obtained from the general model Eq. (5.4), but the fits change little when the submodel Eq. (5.5) is used. The slope estimates from the submodel are nearly identical to those from the general model, and the root mean square error values are similar as well (see Table 5.1).[5]

The plots in Figures 5.6-5.7 display the best-fitting lines obtained from (5.5), along with their fixed points.[6]

These fixed points have abscissa values corresponding to high intensities (105-128 dB SPL) and ordinate values close to zero. The latter indicates that the estimates of $x_*$ and $y_*$ are nearly equal for a given listener and condition. The specific estimates are given in Table 5.1. These values were obtained by least-squares fits of the logarithmic transform of Eq. (5.5) to all the data for a given listener and ISI condition; similar estimates may be obtained from the plots of $10\log K(\nu)$ versus $\beta(\nu)$ in Figure 5.5.

Two listeners were tested under multiple ISI conditions as a preliminary investigation into the possible effect of interval asymmetry on the parameter estimates.

[5] The root mean square error values appearing in the far-right column of Table 5.1 were computed by (i) estimating, for a given listener and ISI, the $\beta(\nu)$ values for all $\nu$, along with the parameters $x_*$ and $y_*$, using the logarithmic transform of Eq. (5.5), and then (ii) fitting the line having slope equal to the $\beta(\nu)$ estimate and intercept equal to $10\log y_* - \beta(\nu)10\log x_*$ to the five data points corresponding to $\nu$.

[6] In presenting Figure 5.3, we are not suggesting that Eq. (5.5) may be extrapolated to describe data obtained for referents of 80 to 130 dB SPL. Indeed, there is evidence that this model would fail at such high intensities (e.g. Viemeister and Bacon, 1988).

The results were not particularly illuminating for the parameters $x_*$ and $y_*$. The variation of ISI did have an impact on the estimated values of these parameters, but for one listener (CD), the relationship was nonmonotonic, and for the other listener (LX), monotonicity was perfectly confounded with the ISI presentation order (first 1000 ms, then 307 ms, then 100 ms). The variation of the $\beta(\nu)$ and $K(\nu)$ estimates with ISI will be examined in future work. For the present work, it is important to note that there was a strong covariation between the $\beta(\nu)$ and $K(\nu)$ estimates for all ISI conditions, supporting the generality of the submodel given by Eq. (5.5).

## 5.4 Discussion

It was argued in this paper that the value of the exponent in the power law modeling the near-miss depends on the choice of discrimination criterion. First, a mathematical result by Falmagne (1985, 1994) was recalled. The result states that if the exponent is different from one for at least one value of the criterion, then the exponent cannot be constant under changes in the criterion. This result—that the value of the near-miss exponent depends on the definition of 'just-noticeable' in the estimation of $x + \Delta(x)$—may be a caution against regarding the exponent as a critical aspect of neural coding of acoustic intensity (cf. Falmagne, 1985).

Next, intensity discrimination data which demonstrate the predicted nonconstancy of the exponent were presented. For criteria ranging from 0.16 to 0.84 and referent levels ranging from 40 to 80 dB SPL, it was observed that a power law provides a good description of the data and that the estimates of the exponent clearly decrease as the criterion increases. Moreover, a striking covariation of the two parameters in this power law was observed. The parameters co-vary in a way that suggests a submodel which has an important fixed-point property: best-fitting lines (in log-log coordinates like those of Figures 5.1-5.4 and 5.6-5.7) for different criteria meet at a common point. This point has an abscissa which appears to correspond to a high intensity, and an ordinate close to zero.

These fixed point estimates and the form of the model specified by Eq. (5.5) lead to the tempting interpretation that sound intensities are evaluated with respect to a high intensity situated at or near the top of the normal range of hearing. This interpretation is consistent with Parker and Schneider (1994), who propose a subjective 'gain control' mechanism which allows the listener to adjust amplification (or attenuation) in the presence of softer (or louder) sounds for improved discriminability. (See also Schneider and Parker, 1990). The idea of a high-level fixed point is not new, as seen in data reported by Stevens (1974): plots of auditory volume versus sound pressure, with tone frequency as a parameter, converge at an abscissa value of about 140 dB SPL. This value is tentatively interpreted by him as a "practical ceiling on the growth of the auditory experience per se" (p. 162). Though the fixed-point estimates in the present study may not correspond to auditory ceilings, they do give additional evidence for the presence of high-intensity standards in the subjective evaluations of sound intensities. Moreover, the present data point to the possible ubiquity of the fixed-point property, which has been observed in several other modalities and which may have additional applications in audition (Stevens, 1974).

Deviations from Weber's law are known in the psychoacoustics literature to depend on such factors as the nature of the stimulus (frequency, intensity range, presence of noise) and the hearing ability of the listener (normal-hearing versus hearing-impaired; see Schroder et al., 1994). The present results give evidence that some deviations also depend on the empirical interpretation of 'just-noticeable.' There are no doubt other experimental factors which determine the extent of the deviation from Weber's law. We currently are developing quantitative models to parse the effects of several of these factors. Aspects of the present data—especially those which speak to the interplay among the exponent, the criterion, interval asymmetry, and interval bias—aid directly in the development of these models.

## 5.5 Appendix

Table 5.2 shows the schedules for adjusting the level of the comparison tone for each of the twelve adaptive tracks used with each referent level in Experiment 1. (Experiment 2 did not include criteria of 0.21 or 0.79.) For six of the tracks, the comparison tone was in the second interval of the 2IFC task (Type 1 trials). For

Table 5.2: Schedule for the twelve experimental tracks used.

| | Trial Type 1 | | Trial Type 2 | |
|---|---|---|---|---|
| Criterion $\nu$ | $r_x$ | $r_y$ | $r_x$ | $r_y$ |
| 0.16 | 4+ | 1− | 1− | 4+ |
| 0.21 | 3+ | 1− | 1− | 3+ |
| 0.29 | 2+ | 1− | 1− | 2+ |
| 0.71 | 1+ | 2− | 2− | 1+ |
| 0.79 | 1+ | 3− | 3− | 1+ |
| 0.84 | 1+ | 4− | 4− | 1+ |

the other six, the comparison tone was in the first interval (Type 2 trials). The level of the comparison tone was contingent on the sequence of responses for the track. Responses indicating that the first or second interval was judged louder are represented by $r_x$ and $r_y$, respectively. Table entries represent the number of consecutive responses needed to change the level of the comparison tone, with the sign indicating whether the level was increased or decreased. For example, for the track corresponding to Type 1 trials with criterion 0.16, the level was decreased by one step following each $r_y$ response and increased by one step following four consecutive $r_x$ responses.

# Bibliography


J.A. Aczél, F.S. Roberts, and Z. Rosenbaum. On scientific laws without dimensional constants. *Journal of Mathematical Analysis and Applications*, 119: 389–416, 1986.

J.B. Allen and S.T. Neely. Modeling the relation between the intensity just-noticeable difference and loudness for pure tones and wideband noise. *Journal of the Acoustical Society of America*, 102(6):3628–3646, 1997.

J.C. Baird and E. Noma. *Fundamentals of Scaling and Psychophysics*. John Wiley and Sons, New York, 1978.

J.E. Berliner and N.I. Durlach. Intensity perception. IV. Resolution in roving-level discrimination. *Journal of the Acoustical Society of America*, 53(5):1270–1287, 1973.

G. Birkhoff. Rings of sets. *Duke Mathematical Journal*, 3:443–454, 1937.

S. Buus and M. Florentine. Psychometric functions for level discrimination. *Journal of the Acoustical Society of America*, 90(3):1371–1380, 1991.

R.L. Causey. Derived measurement, dimensions, and dimensional analysis. *Philosophy of Science*, 36:252–270, 1969.

J.S. Chipman. Consumption theory without transitive indifference. In J.S. Chipman, L. Hurwicz, M.K. Richter, and H.F. Sonnenschein, editors, *Preferences, utility, and demand*, pages 224–253. Harcourt Brace Jovanovich, 1971.

F.L. Dimmick and R.M. Olson. The intensive difference limen in audition. *Journal of the Acoustical Society of America*, 12:517–525, 1941.

C.W. Doble, J.-Cl. Falmagne, and B. Berg. Recasting (the near-miss to) Weber's law, 2003.

C.W. Doble, J.-Cl. Falmagne, B. Berg, and C. Southworth. Systematic covariation of the parameters in the near-miss to Weber's law, pointing to a new law, 2002, submitted.

J.-P. Doignon, A. Ducamp, and J.-Cl. Falmagne. On realizable biorders and the biorder dimension of a relation. *Journal of Mathematical Psychology*, 28: 73–109, 1984.

J.-P. Doignon and J.-Cl. Falmagne. Well-graded families of relations. *Discrete Mathematics*, 173:35–44, 1997.

J.-P. Doignon and J.-Cl. Falmagne. *Knowledge Spaces*. Springer, Berlin, 1999.

A. Ducamp and J.-Cl. Falmagne. Composite measurement. *Journal of Mathematical Psychology*, 6:359–390, 1969.

B. Dushnik and E.W. Miller. Partially ordered sets. *American Journal of Mathematics*, 63:600–610, 1941.

E.N. Dzhafarov and H. Colonius. Multidimensional Fechnerian scaling: Basics. *Journal of Mathematical Psychology*, 45:670–719, 2001.

A. Einstein. *Relativity: The Special and General Theory*. Crown Publishers, New York, 1961. First published in 1916.

J.-Cl. Falmagne. *Elements of Psychophysical Theory*. Oxford University Press, New York, 1985.

J.-Cl. Falmagne. On the interpretation of the exponent in the 'near-miss-to-Weber's-Law'. *Journal of Mathematical Psychology*, 38:497–503, 1994.

J.-Cl. Falmagne. Stochastic token theory. *Journal of Mathematical Psychology*, 41(2):129–143, 1997.

J.-Cl. Falmagne, B.G. Berg, C. Southworth, and M.L. Morrison. Lawful variation of the exponent in the 'Near-Miss-to-Weber's-Law'. Technical Report 96-28, Institute for Mathematical Behavioral Sciences, University of California Irvine, 1996.

J.-Cl. Falmagne and J.-P. Doignon. Stochastic evolution of rationality. *Theory and Decision*, 43:107–138, 1997.

J.-Cl. Falmagne and L. Narens. Scales and meaningfulness of quantitative laws. *Synthese*, 55:287–325, 1983.

J.-Cl. Falmagne and S. Ovchinnikov. Media theory. *Discrete Applied Mathematics*, 121:83–101, 2002.

G. Fechner. *Elemente der Psychophysik*. 1860. Translated by H.E. Adler and published in 1966 by Holt, Rinehart and Winston, Inc., New York.

P.C. Fishburn. Betweenness, orders and interval graphs. *Journal of Pure and Applied Algebra*, 1(2):159–178, 1971.

P.C. Fishburn. Notes and comments: Semiorders and choice functions. *Econometrica*. 43(5-6):975-977. 1975.

P.C. Fishburn. *Interval orders and interval graphs*. Wiley. New York. 1985.

P.C. Fishburn. Generalizations of semiorders: A review note. *Journal of Mathematical Psychology*. 41:357-366. 1997.

P.C. Fishburn and W.T. Trotter. Split semiorders. *Discrete Mathematics*. 195: 111-126. 1999.

M. Florentine. Level discrimination of tones as a function of duration. *Journal of the Acoustical Society of America*. 79(3):792-798. 1986.

M. Florentine and S. Buus. Excitation pattern model for intensity discrimination. *Journal of the Acoustical Society of America*. 70:1646-1654. 1981.

M. Florentine. S. Buus. and C. Mason. Level discrimination as a function of level for tones from 0.25 to 16 kHz. *Journal of the Acoustical Society of America*. 81(5):1528-1541. 1987.

M. Florentine. C.M. Reed. Rabinowitz W.M.. Braida L.D.. and N.I. Durlach. Intensity perception. XIV. Intensity discrimination in listeners with sensorineural hearing loss. *Journal of the Acoustical Society of America*. 94(5):2575-2586. 1993.

S. Gallégo and C. Micheyl. Intensity discrimination and auditory brainstem responses in cochlear implant and normal-hearing listeners. *Behavioral Neuroscience*. 112(4):793-799. 1998.

J.G. Gimbel and A.N. Trenk. On the weakness of an ordered set. *SIAM Journal of Discrete Mathematics*. 11(4):655-663. 1998.

D.W. Grantham and W.A. Yost. Measures of intensity discrimination. *Journal of the Acoustical Society of America*. 72(2):406-410. 1982.

D.M. Green. J. Nachmias. Kearney. J.K.. and L.A. Jeffress. Intensity discrimination with gated and continuous sinusoids. *Journal of the Acoustical Society of America*. 66(4):1051-1056. 1979.

J.P. Guilford. A generalized psychophysical law. *Psychological Review*. 39:73-85. 1932.

T.E. Hanna. S.M. von Gierke. and D.M. Green. Detection and intensity discrimination of a sinusoid. *Journal of the Acoustical Society of America*. 80(5): 1335-1340. 1986.

A. Heathcote, S. Brown, and D.J.K. Mewhort. The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin and Review*, 7 (2):185–207, 2000.

W.S. Hellman and R.P. Hellman. Intensity discrimination as the driving force for loudness. Application to pure tones in quiet. *Journal of the Acoustical Society of America*, 87(3):1255–1265, 1990.

Å. Hellström. Factors producing and not producing time errors: An experiment with loudness comparisons. *Perception & Psychophysics*, 23(5):433–444, 1978.

Å. Hellström. Time errors and differential sensation weighting. *Journal of Experimental Psychology: Human Perception and Performance*, 5(3):460–477, 1979.

C.F. Hix and R.P. Alley. *Physical Laws and Effects*. John Wiley and Sons, New York, 1958.

C.I. Hovland. A note on Guilford's generalized psychophysical law. *Psychological Review*, 45:430–434, 1938.

Y.-F. Hsu, J.-Cl. Falmagne, and M. Regenwetter. The tuning in-and-out model: a random walk and its application to presidential election surveys, to be submitted.

W. Jesteadt, C.C. Wier, and D.M. Green. Intensity discrimination as a function of frequency and sensation level. *Journal of the Acoustical Society of America*, 61(1):169–177, 1977.

S. Kim. On the possible scientific laws. *Mathematical Social Sciences*, 20:19–36, 1990.

D.H. Krantz, R.D. Luce, P. Suppes, and A. Tversky. *Foundations of Measurement (Volume 1)*. Academic Press, New York, 1971.

H. Levitt. Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, 49(2):467–477, 1971.

G.R. Long and J.K.Jr. Cullen. Intensity difference limens at high frequencies. *Journal of the Acoustical Society of America*, 78(2):507–513, 1985.

R.D. Luce. Semiorders and a theory of utility discrimination. *Econometrica*, 24: 178–191, 1956.

R.D. Luce. On the possible psychophysical laws. *Psychological Review*, 66(2): 81–95, 1959.

R.D. Luce. A generalization of a theorem of dimensional analysis. *Journal of Mathematical Psychology*, 1:278–284, 1964.

R.D. Luce. Dimensionally invariant numerical laws correspond to meaningful quantitative relations. *Philosophy of Science*. 45:1–16. 1978.

R.D. Luce. "On the posible psychophysical laws" revisited: Remarks on cross-modal matching. *Psychological Review*. 97(1):66–77. 1990.

R.D. Luce and E. Galanter. Discrimination. In R.D. Luce. R.R. Bush. and E. Galanter. editors. *Handbook of Mathematical Psychology*. volume 1. Wiley. New York. 1963.

R.D. Luce. D.H. Krantz. P. Suppes. and A. Tversky. *Foundations of Measurement III: Representation. Axiomatiation. and Invariance*. Academic Press. San Diego. 1990.

R.J.W. Mansfield. Visual adaptation: Retinal transduction. brightness and sensitivity. *Vision Research*. 16(2):679–690. 1976.

W.J. McGill and J.P. Goldberg. Pure-tone intensity discrimination and energy detection. *Journal of the Acoustical Society of America*. 44:576–581. 1968a.

W.J. McGill and J.P. Goldberg. A study of the near-miss involving Weber's Law and pure-tone intensity discrimination. *Perception & Psychophysics*. 4 (2):105–109. 1968b.

B. Monjardet. Axiomatiques et propriétés des quasi-ordres. *Mathématique et Sciences Humaines*. 63:51–82. 1978.

B.C.J. Moore and D.H. Raab. Pure-tone intensity discrimination: some experiments relating to the "near-miss" to Weber's Law. *Journal of the Acoustical Society of America*. 55(5):1049–1054. 1974.

J.R. Munkres. *Topology: A First Course*. Prentice-Hall. Englewood Cliffs. 1975.

L. Narens. The measurement theory of dense threshold structures. *Journal of Mathematical Psychology*. 38:301–321. 1994.

L. Narens. *Theories of Meaningfulness*. Lawrence Erlbaum Associates. New Jersey and London. 2002.

L. Narens and R. Mausfeld. On the relationship of the psychological and the physical in psychophysics. *Psychological Review*. 99(3):467–479. 1992.

D.L. Neff and W. Jesteadt. Intensity discrimination in the presence of random-frequency. multicomponent maskers and broadband noise. *Journal of the Acoustical Society of America*. 100(4. part 1):2289–2298. 1996.

D.A. Nelson. J.L. Schmitz. G.S. Donaldson. N.F. Viemeister. and E. Javel. Intensity discrimination as a function of stimulus level with electric stimulation. *Journal of the Acoustical Society of America*. 100(4):2393–2414. 1996.

D.K. Osborne. Further extensions of a theorem of dimensional analysis. *Journal of Mathematical Psychology*, 7:236–242, 1970.

E. Osman, H. Tzuo, and P.L. Tzuo. Weber's law, the "near miss," and binaural detection. *Perception and Psychophysics*, 28(4):354–364, 1980.

S. Ovchinnikov. Preference structures I: distances between transitive preference relations. *Journal of Mathematical Sociology*, 3:49–67, 1973.

S. Ovchinnikov. Convex geometry and group choice. *Mathematical Social Sciences*, 5:1–16, 1983.

E. Ozimek and J.J. Zwislocki. Relationships of intensity discrimination to sensation and loudness levels: Dependence on sound frequency. *Journal of the Acoustical Society of America*, 100(5):3304–3320, 1996.

S. Parker and B. Schneider. The stimulus range effect: Evidence for top-down control of sensory intensity in audition. *Perception and Psychophysics*, 56(1): 1–11, 1994.

M.J. Penner, B. Leshowitz, E. Cudahy, and G. Ricard. Intensity discrimination for pulsed sinusoids of various frequencies. *Perception & Psychophysics*, 15(3): 568–570, 1974.

W.M. Rabinowitz, J.S. Lim, L.D. Braida, and N.I. Durlach. Intensity perception. VI. Summary of recent data on deviations from Weber's law for 1000-Hz tone pulses. *Journal of the Acoustical Society of America*, 59(6):1506–1509, 1976.

M. Regenwetter, J.-Cl. Falmagne, and B. Grofman. A stochastic model of preference change and its application to 1992 presidential election panel data. *Psychological Review*, 106(2):362–384, 1999.

R.R. Riesz. Differential intensity sensitivity of the ear for pure tones. *Physical Review*, 31:867–875, 1928.

F.S. Roberts. *Measurement Theory*. Addison-Wesley, Reading, MA, 1979.

F.S. Roberts. Applications of the theory of meaningfulness to psychology. *Journal of Mathematical Psychology*, 29(3):311–332, 1985.

P.N. Schacknow and D.H. Raab. Intensity discrimination of tone bursts and the form of the Weber function. *Perception & Psychophysics*, 14(3):449–450, 1973.

B. Scharf and S. Buus. Audition I. In K. Boff, L. Kaufman, and J.P. Thomas, editors, *Handbook of Perception and Human Performance*. John Wiley and Sons, 1986.

B. Schneider and S. Parker. Does stimulus context affect loudness or only loudness judgments? *Perception and Psychophysics*, 48(5):409–418, 1990.

A.C. Schroder, N.F. Viemeister, and D.A. Nelson. Intensity discrimination in normal-hearing and hearing-impaired listeners. *Journal of the Acoustical Society of America*, 96(5, part 1):2683–2693, 1994.

M. Skandera. A characterization of (3 + 1)-free posets. Unpublished Manuscript, 2000.

J.C. Stevens. *Sensation and Measurement*, chapter Families of converging power functions in psychophysics, pages 157–165. D. Reidel Publishing Company, 1974.

S.S. Stevens. *Handbook of Experimental Psychology*, chapter Mathematics, measurement, and psychophysics. Wiley, New York, 1951.

P. Suppes and J. Zinnes. Basic measurement theory. In R.D. Luce, R.R. Bush, and E. Galanter, editors, *Handbook of Mathematical Psychology*, volume 1. Wiley, New York, 1963.

A.N. Trenk. On k-weak orders: Recognition and a tolerance result. *Discrete Mathematics*, 181:223–237, 1998.

W.T. Trotter. *Combinatorics and Partially Ordered Sets: Dimension Theory*. The Johns Hopkins University Press, Baltimore and London, 1992.

N.F. Viemeister. Intensity discrimination of pulsed sinusoids: The effects of filtered noise. *Journal of the Acoustical Society of America*, 51(4, Part 2): 1265–1269, 1972.

N.F. Viemeister and S.P. Bacon. Intensity discrimination, increment detection, and magnitude estimation for 1-kHz tones. *Journal of the Acoustical Society of America*, 84(1):172–178, 1988.

F.-G. Zeng. Interactions of forward and simultaneous masking in intensity discrimination. *Journal of the Acoustical Society of America*, 103(4):2021–2030, 1998.