**Pacific Northwest
National Laboratory**
Operated by Battelle for the
U.S. Department of Energy

Modeling Power Systems as Complex Adaptive Systems

D. P. Chassin   N. Lu
J. M. Malard   S. Katipamula
C. Posse   J. V. Mallow[a]
A. Gangopadhyaya[a]

[a] Loyola University, Chicago, IL.

December 2004

Modeling Power Systems as Complex Adaptive Systems

D. P. Chassin   N. Lu
J. M. Malard   S. Katipamula
C. Posse   J. V. Mallow[a]
A. Gangopadhyaya[a]

December 2004

Pacific Northwest National Laboratory
Richland, Washington 99352

(a) Loyola University, Chicago, IL.

Abstract

Physical analogs have shown considerable promise for understanding the behavior of complex adaptive systems, including macroeconomics, biological systems, social networks, and electric power markets. Many of today's most challenging technical and policy questions can be reduced to a distributed economic control problem. Indeed, economically based control of large-scale systems is founded on the conjecture that the price-based regulation (e.g., auctions, markets) results in an optimal allocation of resources and emergent optimal system control. This report explores the state-of-the-art physical analogs for understanding the behavior of some econophysical systems and deriving stable and robust control strategies for using them. We review and discuss applications of some analytic methods based on a thermodynamic metaphor, according to which the interplay between system entropy and conservation laws gives rise to intuitive and governing global properties of complex systems that cannot be otherwise understood. We apply these methods to the question of how power markets can be expected to behave under a variety of conditions.

Contents

Figures

Tables

1.0   Executive Summary

The transactive control system is a prototype model of a system of abstract machines interacting using contract networks, a method of distributed system control developed at Stanford University by Reid Smith in 1980. We believe a new theory of control is required to describe the properties of systems of economically and physically interacting devices under this method control. Transactive machines are designed to obey certain local rules already extant in most complex engineered systems. What remains unresolved is a theory using these rules to predict the global average behavior of networked systems composed of large numbers of such machines.

The analogue to such a theory has long existed: statistical mechanics of microscopic atoms and molecules predicts aggregate thermodynamic properties of bulk systems (e.g., pressure, temperature, internal energy). Two requirements of actual machines strongly suggest that ultimately discrete (*viz.* quantum) statistical mechanics may be an appropriate tool: 1) the majority of machines undergo discrete state transitions (e.g., on or off, two-stage, digital control), and 2) some machines have exclusionary properties for certain states which prevent other machines from occupying those states once they occupy them (e.g., ancillary service contracts, remedial action schemes for under-frequency load shedding). Ultimately we seek a model predicting time-dependent behavior of systems of these machines. However, many average properties of these complex systems are time-independent when the system is at equilibrium and can be satisfactorily described using statistical mechanics in the classical limit (for large numbers of machines). Thus we will generally restrict ourselves to modeling time-independent solutions for the scope of this work. With a few exception noted in the text, we leave the treatment of the time-dependent properties of non-equilibrium and properties of small systems to future investigations.

Many of the problems which we seek to solve with respect to power-markets and demand responsive loads ultimately require us to answer a single key question: "what is the probability of a system load Q between the time t and the time $t+dt$?" This question, and a closely related question regarding the price of power, has traditionally been answered using empirical methods. We seek a first-principle method with which we may determine the system load and price of power based on known properties of loads.

In the past, using empirical methods has been an acceptable practice because the load control behaviors were not coupled to the price of power. Therefore feedback mediated by the systems was largely non-existent and the complex emergent behaviors were rarely observed. Indeed, it is the very lack of response to price that has led to current situation in which price volatility can adversely affect system behavior. However, the PNNL's vision of the future is predicated on a change in the assumption and any empirical model of load behavior is therefore called into question. Not having access to real-world data, we have begun the development of simulations that will permit us to observe hypothesized systems. While this is an important step in the right direction, it is not likely to bring us the tools we need to understand the more subtle and rare emergent behaviors that might occur. A first-principles model of price-mediated system behavior is necessary.

Work completed in recent years at PNNL has suggested that statistical mechanics may provide a mathematical framework to achieve this goal. Data collected from various ISOs in recent years reveal

PJM 1999 Price Model
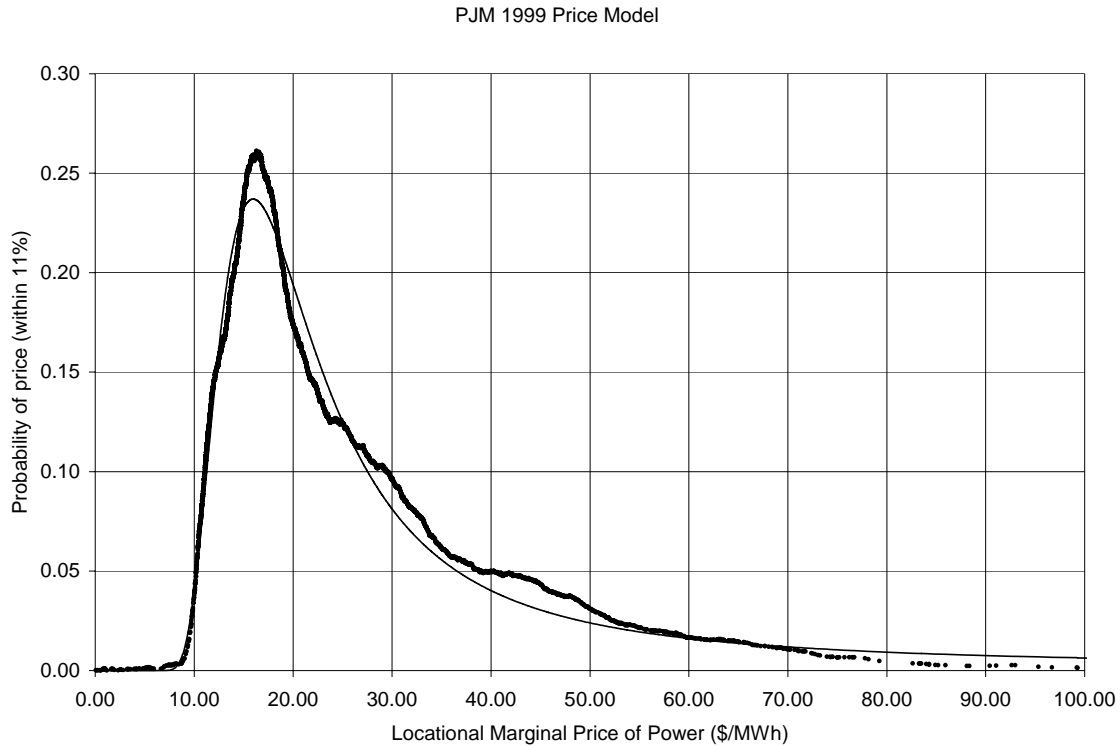
Figure 1: **The Pennsylvania-Jersey-Maryland Interconnect locational marginal price (LPM) history for 1999 (dots) and the theoretical curve (narrow line) for an ideal gas having a Maxwellian velocity distribution as the analog to system load.**

load and price histograms that are reminiscent of Maxwellian probability distribution functions (see Figure 1). Simple models of buyer/seller behavior in power distribution systems have been studied and rigorous definitions for market entropy, value of activity, and market potentials have been derived. Careful analysis of HVAC and waterheater behavior when influenced by market-based prices reveal emergent load behavior that are neither expected, nor allowed for in the grid operations models. We have come to the conclusion that a significant opportunity lies in the application of statistical mechanics to market-based power systems control.

When considering the impact of transactive controls in power transmission and distribution systems, we would like to answer questions such as the following:

- What is the susceptibility of a retail distribution system to changes wholesale market prices? This question goes to the heart of what we mean by demand response. It seems reasonable to expect that distribution systems operated as markets according to the transactive control model will have varying degrees of responsiveness to market price volatility depending on the amount of power trading within them. We would like to quantify this effect for a relatively ideal model such that we can begin to explain the phenomenon more rigorously.

- What is the degree to which markets affect load diversity? We have already seen that markets can decrease load diversity in certain cases. Is it also possible that market mechanics can restore load diversity? We have some evidence to suggest that markets cause "bunching" of the natural

stand-by duty cycles of certain loads. We would like to quantify the effect in a way that allows us to predict whether this is a beneficial and detrimental effect on the power system.

- Certain control strategies (such a frequency-based load shedding) have state-exclusion behaviors that are difficult to model using traditional/empirical load modeling methods. We would like to model these phenomena rigorously enough to predict what level of resources (e.g., Grid-Friendly Appliances) are adequate to provide a certain level of system security.

Modern power system models compute the values for various system properties by explicitly solving the many-body steady–state problem, typically by manipulating a sparse matrix representation of the power flow network topology. However, the introduction of market behaviors and demand response into these systems has created three complications: 1) they have transformed the model into a dense matrix, 2) the dual nature of the model (economic and physical) creates serious obstacles to producing efficient models and solvers, and 3) the scale of the model is increased by several orders of magnitude. As a result, the evolving market-based power grid may not be tractably computable using today's methods and computing systems. Our approach addresses these problems by defining a more suitable model of system properties, which provides useful tools for determine the average properties of systems with extremely large numbers of machines, often many orders of magnitude large than the systems we foresee within the next few decades. Ultimately we expect the mesoscale modeling gap can be addressed by retrodicting the potentials based on the observations of the large scale systems without needing very detailed observation of the microscale behaviors of the individual machines. By recasting the problem in this form, we expect to enable significant contributions from several areas, including 1) q-fermion theory to deduce from the ensemble data the occupation number and set of eigenvalues distribution associated with the underlying quantum potentials (Dutt et al. 1994), 2) solving the inverse scattering problem using super-symmetric quantum mechanics to obtain the set of potentials which produce these eigenvalues (Jeffrey 2004); and 3) applying the various superpotentials from which the family of quantum potentials is derived (Cooper et al. 1994), to produce models for comparison with available data, and to predict the outcome of future market behaviors.

We have observed in data collected that load behavior is not strictly continuous. For example, almost all residential systems have a very small number of discrete states, often only two or three. In addition, the hysteresis of most thermostatic devices (such as heaters, air-conditioners, refrigerators, freezers, and water-tanks) causes them to cycle between two states with a minimum cycle time. Finally, the low degree of precision of typical sensors and the extensive use of digital control cause both the sensing and actuating processes of loads to exhibit discrete state transition phenomenology. Taken together, we conclude that the behavior of a significant number of small machines on power grids is discrete in nature. Therefore we expect that the potentials of the machines, acting as resources in markets would also be discrete in nature. Hence we expect to find a quantum harmonic oscillator, and indeed many other potentials to be relevant to the problems of modeling the interactions of loads in power markets.

2.0   Introduction

The idea of controlling complex systems using decentralized econophysical processes is not new, nor has it always been done with a literal market in the sense that product price and quantity determine the allocation of a scarce resource. Indeed, in the early 1980s, so-called contract networks were demonstrated to allocate scarce central processing unit (CPU) time to competing tasks in high-performance computers (Smith 1980). Nevertheless, the challenges faced by system designers who wish to use economically efficient control strategies are daunting. Such systems are generally called complex adaptive systems because they exhibit a property that is regarded as a profound obstacle to cost-effectively crafting stable and robust classical control strategies: emergent behavior (Kaufman 2000; Gell-Mann 1994; Buchanan 2002; Barabasi 2002). This is the property of systems that causes them to exhibit global behaviors not anticipated by the control strategy. For example, building heating, ventilation, and air-conditioning (HVAC) systems exhibit a phenomenon called global hunting, an accidental artifact of the control strategies used to govern the thermodynamic process (Armstrong 2001). Global hunting exhibits itself as an unexpectedly prolonged cycling between two or more local minima.

Many systems in economics (Smith and Foley 2002; Sergeev 2001; Jaynes 1986), environmental sciences (Georgescu-Roegen 1971), molecular biology (Ptashne 1992), power engineering (Hauer et al. 2001), sociology (Newman 2000), military command/control/communications/intelligence (C^3I) (James and Mabry 2004), and transportation systems, exhibit this unanticipated emergent behavior and are thus candidates for membership in the class of systems characterized by complex adaptive behavior with the potential for unanticipated degradation (Watts and Strogatz 1998; Albert et al. 2000) in robustness. However, exploiting the emergent behavior of complex systems (Feng et al. 2000) by design has been the subject of discussion with respect to building controls (Huberman and Clearwater 1995)and using price as the only signal for electric power systems control (Alvarado 2003). The challenge to exploiting such phenomena is predicting the general characteristics of emergent behavior in complex systems as they relate to the behavior of individual components. This is fraught with analytic difficulties, many of which are discussed by Atmanspacher and Wiedenmann (1999).

To understand the relationship between the rules governing the behavior of system components and the emergent econophysical behaviors of the whole systems, we sought a theory of control based on the thermodynamic analogy (Cont 1999). Tsallis et al. (2003) outlined how to apply nonextensive statistical mechanics to the question of stability and robustness. Donangelo and Sneppen (2003) went further to discuss the dynamics of value exchange in complex systems. In economics, Smith and Foley (2002) show that the correspondence of utility theory to thermodynamics defines a whole consistent methodology, not just a set of analogies. Sergeev (2001) and Jaynes (1986) address the central role of entropy in a thermodynamics-inspired approach to market equilibrium and how it should be used. Jaynes (1991) further expands the analogy for time-varying phenomena using the formalism of predictive statistical mechanics. But Fleay (2001) argues that the conditions imposed by a simplistic thermodynamic approach are utopian and can never be met for real power markets. Nonetheless, the first step in any attempt to understand large-scale system behavior and control is to devise a rigorous, albeit pedagogic, time-independent model of aggregate properties of a system based on properties of the individual components. It is by no means certain that the analogy to statistical mechanics should hold strictly, nor is it necessary. But insofar as it does, we can proceed with the derivation of global properties that are consistent with it.

Pragmatically, we can conclude that the analogy is true if we find no inconsistencies with empirical data or existing theories and if the aggregate properties we develop elucidate the behavior of the systems being considered in a useful way.

2.1   Metaphors for Complex Systems

Physics metaphors have been used repeatedly to model and understand the conceptual and methodological problems arising in other sciences. Classical economics was built largely on the analogy to mechanics. Adam Smith's "invisible hand" (Smith 1776) is a vivid metaphor combining market coordination of selfish individuals scattered throughout a country or even around the world with the idea of an anonymous will pushing toward the social good. This parallels the focus of modern analysis on equilibrium, decentralization and efficiency. This was formalized a century later by Leon Walras in providing the first mathematical statement of general equilibrium theory with rational mechanics.

When Adam Smith wrote The Wealth of Nations, the principles of thermodynamics were not known yet, so the "invisible hand" was interpreted in terms of the era, i.e., a mechanical metaphor. Adam Smith's conceptual model, which identifies the effect of human motivation with forces on the market, gives serious grounds for such an interpretation. In the mechanical models, equilibrium is the state at which the forces applied to the system counterbalance each other and the potential, when steady, reaches its extremum. Consequently, to apply the mechanical metaphor of equilibrium in the economics, some analogs of the mechanical notions are needed. Such a conceptualization is not harmless at all; it implies that the system, having slightly digressed from the state of equilibrium, will return to this very state if left alone. But how this occurs is not stated.

Sergeev (2001) seems to have set out to address this question when he proposed a more general concept of temperature according to which the equilibrium temperature (italics used to indicate an abstract notion as opposed to the physical quantity) of a system is simply the result of the condition that "there are no [net] flows of the conserved macro-parameters between the parts of the system." Thus he defines the notion of temperature as "the derivative of the entropy of the Large System with respect to the macro-parameter for which there exists a conservation law." He makes a particular point to observe that

> *...we said nothing related to either physics, or physical laws and observables.* **All the arguments are applicable to Large Systems of any nature, subject to the above hypotheses**. *These arguments look rather natural, in relation to the large economic systems, if we regard the total income, the total value of products or the total value of consumption of goods as macro-parameters, whereas distribution of income and products of consumption of goods between the subjects of economic activity are viewed as micro-parameters. (emphasis added)*

Sergeev pursues the analogy further to discuss what he terms migration potential, a more economic form of electrochemical potential. He uses it to describe how agents flow from one economic system to another (as opposed to temperature's influence of product flow). He asserts that together these notions are sufficient to describe the equilibrium dynamics of economic systems, stopping short of developing any model for time-dependent dynamics.

Among the valuable contributions made by Sergeev, Smith and Foley (2002) is that according to this method of analysis, one can learn a great deal about the general stability of a complex system by considering the form of the system's entropy function. In particular, entropy functions that have multiple

maxima are characteristic of systems with bifurcation behavior. In addition, certain higher-order entropy functions are characteristic of unstable or runaway systems where in certain circumstances they have no finite maximum entropy. In other circumstances, systems can stagnate when the entropy function is flat within a broad range of states.

Time-independent equilibrium models appear relatively easy to develop, so it is reasonable to expect that the physical metaphor can be applied easily to any complex system—biological, environmental, military, engineering, or other. However, the challenge raised by Fleay (1991) requires consideration of the time-dependent perturbation behavior of these systems. This is a significantly more challenging problem, and it is the ultimate goal of any attempt to model the dynamic behavior of large-scale systems, open or closed.

If there were any lingering doubts about the impact of these ideas on economics, we can hope that Focardi and Fabozzi's (2004) statement will overcome them

> *More recently, the diffusion of electronic transactions has made available a huge amount of empirical data. The availability of this data created the hope that economics could be given a more solid scientific grounding. A new field—*econophysics—*opened with the expectation that the proven methods of the physical sciences could be applied with benefit to economics. It was hypothesized that economics systems could be studied as physical systems with only minimal* a priori *economic assumptions. Classical econometrics is based on a similar approach; but while the scope of classical econometrics is limited to dynamic models of time series, econophysics uses all the tools of statistical physics and complex systems analysis, including the theory of interacting multiagent systems.*

This is as close to an endorsement as is possible today, and it recognizes the potential impact of these ideas on the science of economics and the broader applications to the design and operation of complex economically Pareto-efficient engineered, biological, and environmental systems.

2.2   The Grand Challenge

When adaptable autonomous agents or organisms interact intimately in an environment, such as in predator-prey and parasite-host relationships, they influence each other's evolution. This effect is called co-evolution, and it is the key to understanding how all large-scale complex adaptive systems behave over the long term. Kaufman (2000) suggests that there are two distinct scales of co-evolution: interspecies and system-wide. Interspecies co-evolution is the conventional model and ecological examples abound. Numerous examples exist of avian and insect species with co-evolved feeding and breeding strategies that depend on the parasitic or predatory practices of other species. In almost all these cases, it can be shown that advantage is conferred upon both participants, though it sometimes difficult to discern (as in brood parasites). But it can be argued that the presence of another species or even reciprocal changes in response to it are not sufficient evidence for co-evolution and that there must be evidence from quantitative analysis that shows altered or accelerated evolution of individuals deviated from the expected path based on prevailing conditions. Nonetheless, the conditions for co-evolution arising from cooperative-competitive strategies exist, for instance, in social networks, where criminal informants are used in law-enforcement; economic networks, in stock trading strategies and the Security and Exchange Commission (SEC) rule-making; electric power, by electric power wheeling and Federal Energy Regulatory Commission (FERC) regulatory efforts; military strategy, in arms races.

System co-evolution is a large-scale impact wherein the interaction of one (or a few) co-evolved species with the system as a whole results in changes so fundamental that all species in the system must adapt and the system itself change in significant ways. Certainly, human interaction with the global environment is a commonly cited example of this phenomenon, although we have yet to demonstrate human impact in the genetics of most species. But military planners, economists, engineers, and biologists also encounter such system-wide adaptive "tipping" phenomena in the systems they study.

The phenomenology of co-evolution in biology has been well explored, but what ultimately governs co-evolution from a first-principles perspective is not entirely clear. The question of whether a single ab initio law exists is important to every domain in which complexity is prevalent, whether we consider military strategy, economic policy, global climate change or vaccine development. Certainly entropy maximization and free-energy minimization play important roles in determining which paths a species or a system takes. Morowitz (1992), Bak (1988), and Kaufman (2000) all proposed a so-called "Fourth Law of Thermodynamics" according to which systems tend to self-organize. The requisite conditions vary depending on the author. However, as in the distinction between Clausius (1850) weak statement that entropy *tends* to increase versus Gibbs' (1902) strong statement that entropy *will* increase, we can easily see that any proposed Fourth Law is not a law per se until it can be stated unequivocally. This remains to be done. Indeed, it is not clear that a Fourth Law is necessary if it can be shown that self-organization and co-evolution are a consequence of the balance required by the competing objectives of maximal entropy and minimal conserved quantities. It has not been shown that these are necessary, let alone sufficient. The strongest definition of the Fourth Law is proposed by Liechti: (2000):

> If a system receives a through-flow of exergy (produce*s* entropy/dissipate energy), (a) the system will utilize this exergy flow to move away from thermodynamic equilibrium, (b) if *[…]* more than one pathway to move away is offered from thermodynamic equilibrium, the one yielding most stored exergy, with the most ordered structure and the longest distance to thermodynamic equilibrium, will have a propensity to be selected.

The definition does not match Kaufman's (2000) or Bak's (1988), but it does lead us to consider what happens when several systems interact. As a general approach, we should therefore extend our scope to encompass the interaction of system entropy and the minimization of conserved quantities (selected appropriately according to the domain) as the fundamental insight that must be developed. Unfortunately, this is not likely to be sufficient. The main objection is that it does not anticipate the path-dependent behavior that is often observed in self-organizing systems. However, it is clear that Jaynes (1991) anticipated this when he proposed a generalization of the Carnot cycle as a simple but more useful model to address the problem of explaining path-dependent behavior in complex systems.

The solution to this problem was identified by Smith and Foley (2002) when they proposed using the physical concept of engine cycles. In the context of economic theory, they expressed the motivation for this as follows:

> *Unless the small agents' disequilibrium is constantly replenished, the market maker's activities […] are more relevant to one-shot arbitrage than to the creation of a sustained pattern of activities. A more interesting question is what can be extracted by a small speculator operating between heterogeneous reservoirs, which cannot trade with each other directly. This becomes a problem for the speculator when the good which she can readily exchange (say, shares) is one for which the reservoirs do not have disparate prices, or in which they do not trade at all.*

According to this model, agents capable of storing small amounts of a conserved product interact alternatively with two systems having differing temperatures with respect to the conserved product's parameters. This alternating interaction gives rise to an engine capable of extracting a conserved product from the flow between the two systems mediated by an agent. The total work done is simply the area inscribed by the system's trajectory. By *work* we mean the quantity of product taken from the source system but not delivered to the sink system and thus diverted for some unspecified purpose. Some important observations can be made about systems that behave in this manner. In particular,

1. The process followed by such an engine is not perfectly efficient and introduces entropy during every cycle. As a result, the engine is not fully reversible and the interactions between the two systems are not fully reversible.

2. The definition of work is based on the product transfer process, not on properties of the agent itself. Thus, the difference in any two states of the systems depends on the path taken and not on the states themselves. Unlike entropy and temperature, the work done is not an intrinsic property of systems. In an economic system, one can think of work as the source of profit.

3. The use of such a model requires us to define something like *free product*—that fraction of the total product available that may be ideally extracted from a system using such an engine. As with the notion of free energy, free product is a very important intrinsic property of any complex system.

4. The efficiency of such an engine is determined exclusively by the difference between the temperatures of the systems it is connecting to, and is not a property of the engine itself. Otherwise, one can show that non-conserving agents could be constructed, which is impossible or at best unsustainable.

5. Such an engine can be run in reverse and can use product to move product in the opposite direction it would naturally follow based on the temperature gradient. In an economic system, one can think of the product used as investment in the system of higher *temperature* and exploitation of the system with lower *temperature*.

The generalization of this model to military, law-enforcement, and engineering systems could be very helpful in understanding some of the most challenging phenomena we observe.

2.3   Biology Challenges

Entropy computations play an important role in understanding the spatial patterns of gene expression, waves of low chemical concentrations in circadian clocks and along life-sustaining chemical pathways. Entropy also plays a critical role in understanding from a statistical perspective how noise in chemical concentrations and locations enhance and amplify the many weak signals that are essential to a healthy cell.

The sequencing of the human genome opens the door for vast numbers of new drug targets. It also increases the complexity of the task because it enables the study of drug interactions. Drug development is an expensive design-oriented process, and methods are needed to select effective drug targets as early as possible. Fuhrman et al. argue that time series from gene expression data with a high entropy are more likely to be effective drug targets. Langmead et al. (2002) also use entropy-based methods (including the

Shannon extension in information processing) to detect periodic expression patterns in gene expression data. Nemenman et al. (2003) further extend this approach to address biological time-series problems:

> *The major problem in information theoretic analysis of neural responses and other biological data is the reliable estimation of entropy—like quantities from small samples. [A recently introduced Bayesian entropy estimator…] performs admirably even very deep in the under-sampled regime, where other techniques fail. This opens new possibilities for the information theoretic analysis of experiments, and may be of general interest as an example of learning from limited data.*

Indeed, timing is critical in many biological processes, but it is measured differently at different scales, ranging from the frequency of chemical waves at the cellular level up to the variations of daylight and ambient temperatures at the organism level. At every level, the clocks must be synchronized. Time can be measured at any level because some change occurred, e.g., in chemical concentrations or in energy flux. Biological macrosystems may therefore have a faster clock cycle than some of their subsystems. Rojdestvenski and Cottam (2000) show how the concept of entropy can be used to estimate the difference of clock speed based on the number of states in the systems. Furthermore, according to Vilar *et al.* (2002), noise can actually enhance the weak signals that drive the clocks:

> *A wide range of organisms use circadian clocks to keep internal sense of daily time and regulate their behavior accordingly. Most of these clocks use intracellular genetic networks based on positive and negative regulatory elements. The integration of these "circuits" at the cellular level imposes strong constraints on their functioning and design. [... This] type of oscillator is driven mainly by two elements: the concentration of a repressor protein and the dynamics of an activator protein forming an inactive complex with the repressor. Thus the clock does not need to rely on mRNA dynamics to oscillate, which makes it especially resistant to fluctuations. Oscillations can be present even when the time average of the number of mRNA molecules goes below one. Under some conditions, this oscillator is not only resistant to but paradoxically also enhanced by the intrinsic biochemical noise.*

In fact, it has been known for a decade that the response of nonlinear systems to weak signals may be enhanced by noise, including animal feeding behavior, human tactile and visual perception, and some neurobiological systems (Yam02 and references therein). Such phenomena are examples of the so-called stochastic resonance statistical dynamics. Perc and Marhl (2004) rigorously studied stochastic resonance effects in the response of intracellular Ca2+ to weak signals. They show how different noise intensities enhance optimally different signal frequencies. Kummer and Ocone (2003) show how the concept of entropy leads to a thermodynamic formulation of the epigenic system in which the pseudo (or talandic) temperature measures the chemical oscillations between genetic locus and the resulting metabolites. Their model is simplified but suggests how entropic processes play a central role in understanding the adaptive behavior of complex macroscopic biological systems.

Another area where the interplay between system entropy and conservation laws may play a major role is cellular morphology, the study of how stem cells either proliferate or generate specialized descendants (Kaneko and Yomo 1997). Silva and Martins (2003) apply entropy related concepts within a cellular automata framework and use a model of gene coupling to study the frequency of specialization.

Another area where the interplay between system entropy and conservation laws may play a major role is cellular morphology, the study of how stem cells either proliferate or generate specialized descendants (Kaneko and Yomo 1997). Silva and Martins (2002) apply entropy-related concepts within a cellular automata framework and use a gene-coupling model to study the frequency of specialization.

These examples illustrate that entropic approaches to modeling complex biological systems may yield new insights and accurate predictions not achievable by entirely atomistic models or macroscale bulk models. Two challenges also emerge. First, biological systems are inherently hierarchical with coupling across levels. Entropy estimates at one level must be related to those at the next level. One illustration of this is that while the fast interactions between small molecules can be accurately modeled using reaction rates, the interactions with and between proteins involves relatively slower conformational changes and need not be accurately modeled using simple reaction rates (Kurzynski 2000). Deterministic models of multiscale systems face the same challenge; however, in the present case the coupling is empirical. Cross-design for very large statistical linear models with 100,000,000 equations is being investigated by Abowd (2003). Thus, the first hurdle may not be so daunting.

A second challenge is that most biological processes appear discontinuous at a small time scale. The last 100 years have seen development of mathematical analytical techniques that address specifically periodic phenomena; most biological processes in a healthy organism are cyclic. Drawing on the experience gained from mathematical physics, basic research is needed to build accurate quantitative models of realistic biological systems. As Patel et al. (2000) put it in the context of a quantum DNA search problem,

> *Identification of a base-pairing with a quantum query gives a natural (and first ever) explanation of why living organisms have 4 nucleotide bases and 20 amino acids. It is amazing that these numbers arise as solutions to an optimisation problem.*

2.4 National Security Challenge

Small world models of social networks have produced a great deal of literature on the subject of punctuated equilibrium dynamics. Bak and Sneppen (1993) first brought together the concept of punctuated equilibria and self-organizing criticality in addressing complex evolutionary systems. The parallel between species in an ecological environment and clusters in a social network seems to suggest itself with little effort.

Recent work comparing the properties of social clusters to those of spin clusters (Shafee 2004) shows parallels between phase changes in market and labor stratifications and determined the requisites for redrawing the membership boundaries with respect to energy minimization and efficiency. This work has shed some light on quantifiable variables relating to the concept of "self" and its importance in the evolution of social clusters. Distorted ideas of self can account for altruism in such systems, and by extension one expects them to account for other forms of extreme individual behavior, including sociopathy.

The infrequent upsurge of anomalous and extreme social behavior can also be approached by studying phenomena such as the onset of synchronization in networks of coupled oscillators (Moreno and Pacheco 2004). Averaging theory has been used to determine the ranges of parameters for which network

oscillations are damped as the system approaches a global steady state (Atay 2004). The application of this to any periodic process and in particular to social phenomena is readily apparent.

By simple extension of the Bak–Sneppen (1993) model, we can see that the stability of a social cluster is determined by the barrier height separating its local fitness maximum from other better maxima. In the case of an individual in a social network, the barrier height might simply be the number of beliefs that must be changed. Single belief changes occur often, but complicated changes such as radical multidimensional belief revisions are prohibitively unlikely because they involve complex and coordinated moves through an ill-defined landscape. Thus, the time scale for change is exponential in the barrier height. When the fitness of belief is high, it is often very difficult to find a nearby maximum that is better. Those states are relatively stable. When the fitness is low, it is very likely that a nearby better state can be found, so the barriers are low. The challenge for modelers of social networks in the context of sociopathic behavior is that the ability of individuals to quickly find erroneous maxima has been greatly increased by the widespread availability of dubious information and the increasing resonance of provocative messages.

Without much comment on its obvious historical impact, Carvalho-Rodriguez et al. (1993) studied the loss of social cohesion resulting from pandemics between the 14th and 17th centuries in Europe. Based on the data, they suggest that a critical threshold exists near 37% mortality, at which point the disruption of the social structure is maximal. But they also suggest using this entropy measure to gain insight into the impact of low-intensity conflict on the stability of civil government.

According to Herman (1999), conflict models are fundamentally attrition-based, and entropy is the macroscopic expression of the combined effect of Clausewitzian friction, disruption, and lethality. Taken from another perspective, entropy collectively expresses unit cohesion and capability—as a unit loses cohesion, its entropy rises and its capabilities decrease. Combat casualty models have also been developed based on entropic processes. Dexter (2003) investigated casualty based entropy and the entropic phase space developed by Carvalho-Rodrigues et al. (1992) to determine their usefulness as measures of combat effectiveness. Dexter extended the combat entropy method (an entropy difference) for time-independent data to show its utility for time-dependent data, where Carvalho-Rodrigues applied an entropic phase space. While the entropic phase space was determined to be a poor outcome predictor, Dexter found that the combat entropy between two forces is a good predictor of attrition-based outcomes wherein the successful force is usually one that can maintain lower entropy for a greater time period.

Models such as this show a great deal of promise for detecting and monitoring anomalous social and organizational behavior. However, we are a long way from implementing these in practical systems. Furthermore, visualization of social network analysis systems presents an awesome challenge that has yet to be addressed satisfactorily. Reviewers of the subject often suggest that visualization is critical to the progress of scientific fields (Freeman 2000). Freeman discusses at length how computing technology has affected the study of social systems, including procedural analysis methods that use the interaction of viewers with network data to explore their structural properties. This approach has taken root in the work done by the leading visualization research groups around the world, including those at research laboratories of the Department of Energy, National Institutes of Health, Department of Commerce, Department of Defense, and various law enforcement agencies. It is clear that visualization technology is an essential element to moving the existing theoretical body of knowledge about the behavior of complex adaptive systems into practical domain-specific and application-oriented technology suites that can address the most pressing problems of the day.

2.5 Engineering Challenge

System complexity is a long-term challenge to the design and operation of many engineered systems, particularly those that interact with market-based financial systems such as the transportation logistics network, telecommunications networks, and electric and gas energy delivery networks. Developing tools and techniques to manage system complexity is essential to 1) developing an understanding of how technologies, policies, and regulations will affect the stability of complex, evolutionary, economically driven systems and 2) monitoring markets in real time to ensure they are not being manipulated by the invisible hand of a cheater who violates the rules.

The phenomena that must be studied comprise multiple independent devices and agents operating at a variety of scales—physical scales ranging from continents down to devices within individual homes and businesses and time scales ranging from milliseconds to decades. For example, when the electric grid actively engages end-use appliances and equipment as integral elements of system control, the scale of the U.S. system expands from roughly 100 thousand electrical busses to more than 1 billion pieces of equipment, a 10-thousand-fold growth. Transient stability phenomena are driven by subsecond electromechanical and control processes, while planning and construction of power plants, transmission lines, and distribution substations takes place over the course of years and their economic lifetimes extend over decades. The ability to understand and manage these phenomena involves answering a variety of fundamental scientific challenges.

Recent research by Oliveira et al. (2003a) as extended mathematical methods originally developed for exploring the complex combinatorial problem of ultra-large biological networks (Oliveira et al. 2003b) to the problems posed by the power grid. Known as Petri networks, this method relies on a decision net-work that uses mathematical equations to express complex and highly dimensional relationships and can be used to construct models. Petri nets have been used efficiently and comprehensively to precompute the boundaries of network stability regions, and the inter-network influence relationships, reducing the computational complexity associated with this class of network-of-networks that are hyper-exponentially hard to solve to near linear time complexity. Most importantly they have shown the ability to identify system invariants that correspond to conservation laws in the traditional engineering models.

Additional research has been conducted into an underlying theoretical basis for the behavior of most market-governed large-scale engineering system. Development of this theory of transactive control can be used to predict and manage complex engineering system behavior (Chassin 2004a). Similar to the way statistical thermodynamics is used to describe large numbers of molecular particles, the theory of transactive elements applies statistical methods to analyze the overall functional properties of populations of machines as they obey conservation laws for energy trading in a transactive energy grid. This approach already has revealed profound misconceptions about the behavior of loads in the presence of utility demand response programs (Lu and Chassin 2004). It also has been shown to predict distributions of electricity prices that match observations from some of today's wholesale markets (Chassin 2003). It is hoped that such a theory will eventually provide the basis for knowing whether a market is operating in an abnormal condition (i.e., being manipulated).

Additional research has been conducted into an underlying theoretical basis for the behavior of most market-governed large-scale engineering system. Development of this theory of transactive control can be used to predict and manage complex engineering system behavior (Chassin 2004a). Similar to the way statistical thermodynamics is used to describe large numbers of molecular particles, the theory of

transactive elements applies statistical methods to analyze the overall functional properties of populations of machines as they obey conservation laws for energy trading in a transactive energy grid. This approach already has revealed profound misconceptions about the behavior of loads in the presence of utility demand response programs (Lu and Chassin 2004). It also has been shown to predict distributions of electricity prices that match observations from some of today's wholesale markets (Chassin 2003). It is hoped that such a theory will eventually provide the basis for knowing whether a market is operating in an abnormal condition (i.e., being manipulated).

Ultimately, we need modeling approaches that allow for evolutionary changes to the strategies of behavioral agents to develop the decision support tools for policy makers and regulators who are concerned with preserving public and private benefits. The traditional deterministic approach used by engineering simulation is woefully inadequate, hence the need to understand better the role of entropic equilibrium and its relationship to system conservation laws. Experimental economics has given us some insights into what these conservation laws might be and how individual decision processes contribution to the entropy of engineered econophysical systems. Vernon Smith, winner of 2003 Nobel Prize for economics, pioneered techniques of experimental economics that used people as actors in a gaming environment to explore the distinction between actual behavior and theoretically ideal behavior. HyungSeon et al. (in Stoft 2002) reported being successful at developing insights into the strategies and differences between automated and human agents as they operate in power markets. The next step is to incorporate the exploratory behavior of humans into the engineering models of system behavior such that the economic motives of player-agents are more accurately reflected in the long-term dynamics of the system. This will enable the study market structures and reveal how economic and technical control strategies co-evolve over time so that appropriate rules and policies for regulating the national infrastructures can be comprehensively explored.

3.0   Models

To understand the relationship between the rules governing the behavior of devices and the emergent behaviors of the systems, we established a theory of control that applies to complex adaptive systems. While the full development of such a theory is beyond the scope of this report, it is necessary to devise a rigorous model for the constituent devices in such systems. We expect that such a rigorous model will be able to postulate laws and ultimately develop definitions of average properties of complex systems that will be experimentally verifiable.

3.1   Abstract Machine Model

The first step is to establish the basic concepts of the complex adaptive systems to be considered. First we define an abstract machine,[a] the basis for all devices we can conceive in real systems. Figure 3.1 illustrates the abstract machine used in all considerations. This abstract transactive machine is the basis for coupling a thermodynamic work process to an economic control process. The specifics of these two processes will be left to future discussions on the development of rules and the resulting system behavior. The abstract transactive machine converts in time t a quantity Q_{int} of resources obtained from another transactive machine at cost C_{outt} and produces quantity Q_{outt} of resource sent to another transactive machine for which it receives compensation, Cint.

> *A transactive machine is a device that converts in a time* t *a quantity* $Q_{in}t$ *of resources obtained from another transactive machine at a cost* $C_{out}t$ *and produces a quantity* $Q_{out}t$ *of resource sent to another transactive machine for which it receives compensation* $C_{in}t$.

The device may also produce a waste at a rate Q_{waste}, for which no consideration can be derived, and the device may require the investment of capital at a cost $C_{invest}t$ or produce $C_{profit}t$, all of which play a role in the basis for the device's value. A number of variations on this model are possible, but fundamentally every conceivable engineered device operating in an economic system should be based on this abstract design.

The device must be a deterministic machine in the sense that at any given instant, if a state change in the thermodynamic process occurs that is communicated to the control logic, the control logic will select a new state for the machine in no greater than polynomial time with respect to the number of possible states sensed.[a] Therefore, given sufficient settling time, the device will always be in a single known state chosen from among a finite and unchanging set of states. The term unchanging means that the number of accessible states does not change perceptibly in the time scale in which a device's state may be observed. The set of possible states remains unchanging for the duration of observations and actions of the control logic.

(a)   In the context of this theory the device includes both the machinery that performs the thermodynamic process and the control logic that is called the controller or the agent.

(a)   This is the most important criterion for Turing determinism as it applies in the context of this discussion. If the sensor data from the thermodynamic process has a finite number of states (it does so in digital control systems), the control state sent to the process must be determined by the controller in a time that is a polynomial function of the number of sensor states. This assures us that the device will always be in a correct state or in a transient state en route to a correct state that is reached in a finite time.
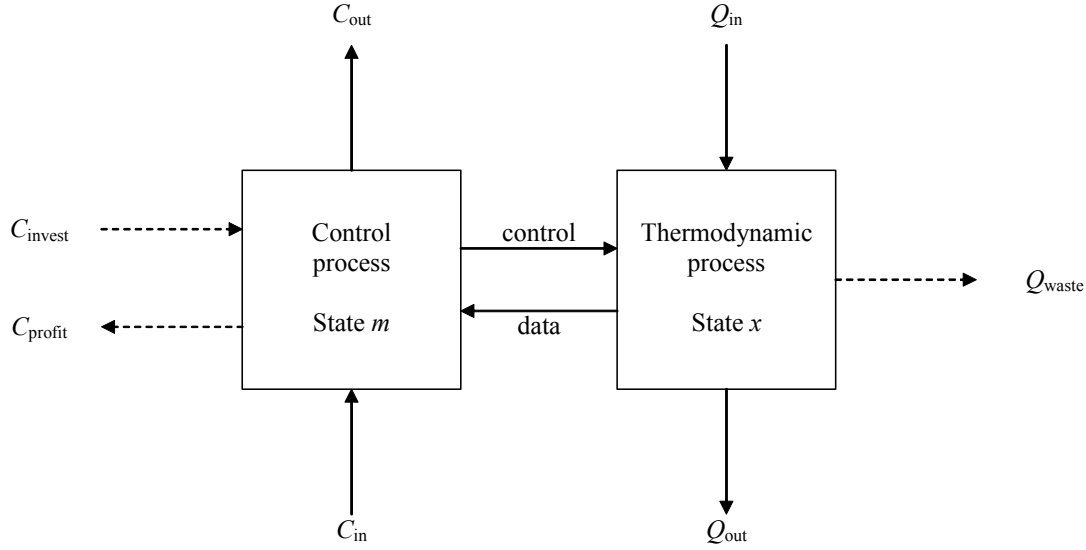
Figure 3.1.   Abstract Transactive Machine

We should also note an important consideration regarding the state of a device. The control logic of the device may in fact completely disconnect the device from a system, i.e., reduce all rates to zero. In another case, devices may be created and destroyed according to rules for their currency balance, removing them from the system. In this sense, we must use caution when considering the number of devices in a system. The number of devices may not be constant depending on the characteristics of the devices involved. To use an analogy to thermodynamics, an ideal gas has a constant number of particles and thus equations of state where energy is a linear function of temperature. In contrast, a photon gas has a variable number of particles and the internal energy function is volume dependent. Thus we must consider the possibility that systems in which devices may turn off are distinct from systems in which all devices must always have at least one non-zero rate and a non-zero currency balance. The only case in which such a distinction would be unnecessary would be if the "off" state of a device has a non-zero potential. This would be true if turning a device off had value in a system, such as when under-frequency load shedding is used to maintain frequency stability on power grids (Grigsby 1999).

3.1.1   Periodic Behavior of Abstract Machines

In the vernacular of power systems, the amount of energy consumed by a load is usually measured in watt-hours (W·h). We call this value the *position* of the load. The instantaneous amount of power used is typically measured in watts (W) and is called the *rate* of a load. In theory, a change in the rate of a load is referred to as the ramp rate and is measured in watts-per-hour (W/h). Often this change is an abrupt step-function, so the integral of the change is finite but its slope is infinite. Such a change in the rate of a load is referred to as a *transaction*. The instantaneous nature of some transactions recommends the use of the Dirac delta function, δ,[a] to model transactions. This model, described only in terms of loads, can be

---

(a) The Dirac delta function is really a *functional,* even though that notion hadn't been invented when Dirac first described it. The delta function has three important properties: 1) it is a probability density function, i.e., $\int_{-\infty}^{\infty}\delta(x-a)dx=1$; 2) it has a variance of zero, i.e., $\delta(x-a)=0$ $x \neq a$; and 3) it selects the value of an arbitrary function f evaluated at the location of the point mass, i.e., $\int_{-\infty}^{\infty}f(x)\delta(x-a)dx=f(a)$.

applied as well to any generation source capable of instantaneous change in power output. This section focuses on modeling periodically recurring loads whose transactions are modeled with the delta function.

Therefore, let $\tilde{\delta}$ be the periodic delta function of period ω and phase $0 < \varphi < 2\pi$ such that $\int_{-\pi/\omega}^{\pi/\omega} \tilde{\delta}(\omega t - \varphi) = 1$ and $\int_{0}^{t} \tilde{\delta}(\omega x - \varphi) dx = \mathrm{E}\left(\frac{\omega t - \varphi}{2\pi}\right)$.[b] The positive form of the $\tilde{\delta}$ function defines a periodic load creation transaction, given by $T_{+}(t) \equiv \tilde{\delta}(\omega t)$, and the negative form defines the periodic load annihilation transaction, $T_{-}(t) \equiv -\tilde{\delta}(\omega t - \varphi)$.

Consider the limit of the integral of the sum of both transactions as time approaches infinity. We find that

$$\lim_{t\to\infty} \int_{0}^{t} \left[T_{+}(t) + T_{-}(t)\right] dt = \lim_{t\to\infty} \int_{0}^{t} T_{+}(t) dt + \lim_{t\to\infty} \int_{0}^{t} T_{-}(t) dt$$

$$= \lim_{t\to\infty} \left[\mathrm{E}\left(\frac{\omega t}{2\pi}\right) - \mathrm{E}\left(\frac{\omega t - \varphi}{2\pi}\right) \right] \qquad (3.1)$$

converges and

$$\lim_{t\to\infty} \int_{0}^{t} \left(T_{+}(t) + T_{-}(t) \right) dt = \frac{\omega \varphi}{2\pi} \qquad (3.2)$$

As shown in Figure 3.2, the position X of the load is the integral over t of rate Q, so

$$Q(t) = X'(t) \qquad (3.3)$$

and the transaction T is the derivative of Q with respect to t, so

$$T(t) = Q'(t) = X''(t). \qquad (3.4)$$

When $\omega(b-a) \gg 2\pi$, we can approximate $\int_{a}^{b} \tilde{\delta}(\omega t - \varphi) dt \approx \omega \frac{b-a}{2\pi}$. When $t \gg 1/\omega$, we can represent the average rate, Q, of a periodic binary load between time 0 and time t by

---

(b) The function E(x) finds the nearest integer $n \leq x$ such that $x - n < 1$.

$X(t) = \int Q(t)$

$2\pi/\omega \quad \varphi$

$Q(t) = \int T(t)$

$T_+(t) = \delta(\omega t) \qquad t \qquad\qquad t+2\pi/\omega \qquad t+4\pi/\omega$

$T(t) = T_+(t) + T_-(t)$

$T_-(t) = -\delta(\omega t - \varphi) \qquad t+\varphi \qquad t+2\pi/\omega+\varphi \qquad t+4\pi/\omega+\varphi$
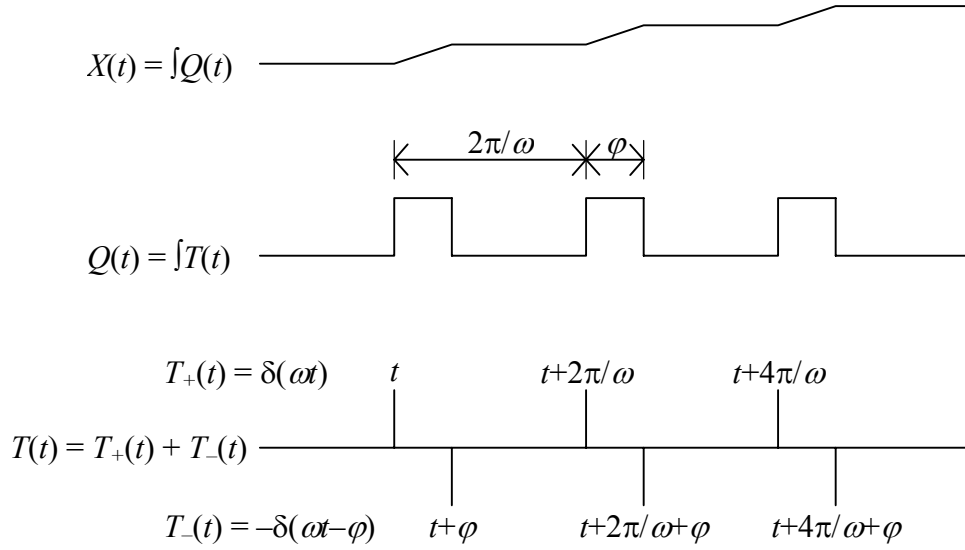
Figure 3.2.   Position, Rate, and Transaction

$$\langle Q(t) \rangle = h\left(\int_0^t \tilde{\delta}(\omega x)dx - \int_0^t \tilde{\delta}(\omega x - \varphi)dx \right) = h\left(\omega \frac{t}{2\pi} - \omega \frac{t-\varphi}{2\pi} \right) = \frac{h}{2\pi}\omega\varphi \qquad (3.5)$$

where *h* is the size of the transaction defined as a unit of load for a unit of time. If arbitrarily we take $h = 1$ Wh^2 (1 Wh load contract for 1 hour), we have $\hbar = {}^1\!/_{2\pi}$ (the number of *h*'s per 2π hours, measured in Wh^2) and

$$\langle Q(t) \rangle = \hbar\omega\varphi \qquad (3.6)$$

We must note that $Q(t) = h$ during the active phase, and $Q(t) = 0$ during the inactive phase, so the quantity $\omega\varphi/2\pi$ is the probability of the load *h* at any given time. Therefore,

$$\langle Q(t)^2 \rangle = h^2 \frac{\omega\varphi}{2\pi} = 2\pi\hbar^2\omega\varphi \qquad (3.7)$$

Thus, while the variance of *T* is zero for all time, the variance of $Q(t)$ is given by

$$\sigma_Q^2(t) = \left(2\pi\hbar^2\omega\varphi\right) - \left(\hbar\omega\varphi\right)^2 = \hbar\omega\varphi\left(2\pi\hbar - 1\right) \qquad (3.8)$$

3.1.2   Hot Water Tank Standby Load

Consider an electric residential hot water tank that is standing idle. It periodically must turn on its heating element in order to maintain the temperature of the water within a certain range. The approximate water temperature at the time *t* is given by the first order differential equation

$$T_w'(t) = \frac{Q(t)}{C} - \frac{HT_w(t)}{C} + \frac{HT_{air}(t)}{C} \qquad (3.9)$$

3.4
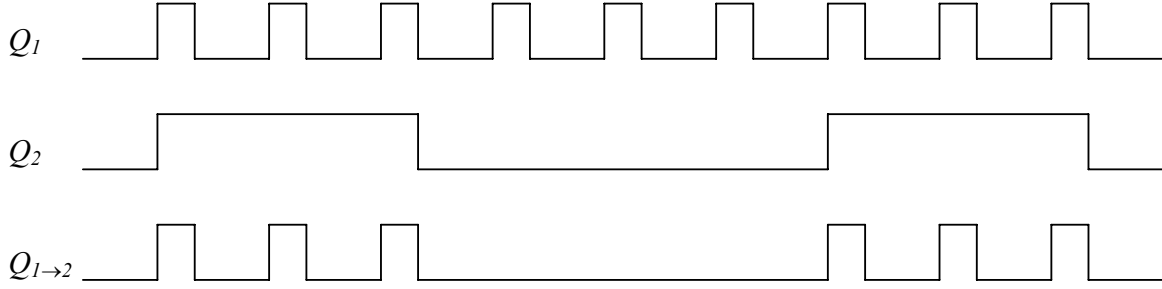
Figure 3.3.   Composite Periodic Loads

where T_w is the water temperature, T_{air} is the ambient air temperature, C is the heat capacity, H is the thermal conductivity of between the water and the ambient air, and Q is the heat input from the heating element.

The heating element is turned on (active $\Rightarrow Q(t) = h$) only when the temperature T_w falls below the T_{on} and turned off (inactive $\Rightarrow Q(t) = 0$) only when it rises above T_{off}.[a] The frequency of the switching regime may be determined by solving Eq. (3.9) for the duration of the active and inactive periods, respectively φ and $\varphi_{\dagger}$, where $\omega(\varphi + \varphi_{\dagger}) = 2\pi$. Thus we find

$$\varphi_{\dagger} = \frac{C}{H}\log\left(\frac{T_{off} - T_{air}}{T_{on} - T_{air}}\right)$$

$$\varphi = \frac{C}{H}\log\left(\frac{HT_{off} - HT_{air} - h}{HT_{on} - HT_{air} - h}\right)$$

From this we can determine the period

$$\omega = \frac{2\pi}{\varphi + \varphi_{\dagger}} = \frac{2\pi H}{C\log\left[\dfrac{H - h\left(T_{off} - T_{air}\right)}{H - h\left(T_{on} - T_{air}\right)}\right]} \tag{3.10}$$

The average value of Q up to the time t is given by Figure 3.3, so the average load L of N independent hot water tanks is given by

$$\langle L(t)\rangle = N\langle Q(t)\rangle = N\hbar\omega\varphi$$

$$= Nh\left(\frac{\log\left[H\left(T_{off} - T_{air}\right) - h\right] - \log\left[H\left(T_{on} - T_{air}\right) - h\right]}{\log\left[H - h\left(T_{on} - T_{air}\right)\right] - \log\left[H - h\left(T_{off} - T_{air}\right)\right]}\right) \tag{3.11}$$

(a) It is impossible to determine the value of $Q(t)$ from T_w when $T_{on} < T_w < T_{off}$ without also knowing T'_w and ensuring that no demand load is lowering T_w. Fortunately, this approach to solving the load does not require that knowledge, in contrast to both finite difference and differential models, which require "stateful" finite state machines to be computable.

3.1.3 Composite Loads

Combining M interdependent loads Q of differing periods produces composite periodic loads, so that any complex load is given by

$$Q(t) = \prod_{i=1}^{M} Q_i(t) = \hbar^{-M} \prod_{i=1}^{M} \frac{\varphi_i}{\omega_i}. \tag{3.12}$$

Thus we find that

$$\log Q_{1 \to M}(t) = \int \frac{dQ_{1 \to M}(t)}{Q_{1 \to M}(t)} = \int \frac{T_{1 \to M}(t)}{Q_{1 \to M}(t)} = \sum_{i=1}^{M} \frac{\varphi_i}{\omega_i} - M \log \hbar = M\left(\overline{\theta} - \log \hbar\right) \tag{3.13}$$

where

$$\overline{\theta} = \frac{\sum_{M} \frac{\varphi_i}{\omega_i}}{M}. \tag{3.14}$$

So we have

$$Q_{1 \to n}(t) = e^{n\left(\overline{\theta} - \log \hbar\right)} \tag{3.15}$$

3.1.4 On the Observability of Load Events

Often it is said that a signal cannot be localized simultaneously in time and frequency. This statement is misleading because the signal itself is not concentrated in time and in frequency. Thus we recognize that the Heisenberg Uncertainty principle, for example, is not a statement of *observability* but of *reality* (Hubbard 1998). The shorter the duration of a load, the wider the band of frequencies given by the Fourier transform; the narrower the band of frequencies of its transform, the more the load is spread in time. This is stated precisely by saying that for every load $Q(t)$ (where t is a real number) such that

$$\int_{-\infty}^{\infty} |Q(t)|^2 \, dt = 1 \tag{3.16}$$

the product of the variance of t and the variance of ω (the variable of $\hat{Q}$) is at least $\frac{1}{16\pi^2}$:

$$\underbrace{\left(\int_{-\infty}^{\infty} (t - \overline{t})^2 |Q(t)|^2 \, dt\right)}_{\text{var}(t)} \underbrace{\left(\int_{-\infty}^{\infty} (\omega - \overline{\omega})^2 |\hat{Q}(\omega)|^2 \, d\omega\right)}_{\text{var}(\tau)} \geq \frac{1}{16\pi^2} \tag{3.17}$$

(The exact value of the minimum on the right depends on the formula used for the Fourier transform, which affects where the 2π is placed. The normalizing condition is used so that this integral measures a probability of 1, which describes an event that is certain.)

These variances measure the extent to which t and ω take values far from their respective means, $\bar{t}$ and $\bar{\omega}$. Thus the more Q is concentrated in a small window of time, the smaller the variance of t will be; if the signal is more spread out in time, the variance will be larger. Multiplying $|Q(t)|^2$ by $(t - \bar{t})^2$ emphasizes the values of $|Q(t)|^2$ for larger values of $|t - \bar{t}|$ when t is far from its average and minimizes them when $|t - \bar{t}|$ is small. The second integral in Eq. (3.17) measures the range of frequencies by the transform of Q. The smaller the variance ω, the narrower the band of frequencies of Q.

So consider a load $|Q(t)|^2$ whose graph is very concentrated around $t = 0$. Normalizing the integral of this load at 1, the load looks like a peak and the Fourier transform of a peak is necessarily very spread out. Therefore, whatever method we use to decompose the signal into time and frequency, we will run into this problem. Precise information about the time of a load event comes at the price of vagueness about the frequency with which it can occur. Precise information about frequency load events comes at the price of vagueness about when they occur.

3.1.5 Are Power Markets Quantum Markets

It is known that power markets may have no clearing price at which supply is equal to demand. We explore this phenomenon to understand how a quantum model may be a more accurate model with which to design power markets that must consider the role of demand response. The advantage of a quantum model is that it not only provides the basis for demonstrating how demand response increases systems efficiency and robustness, but it also give us a clearer picture of how these advantages might be achieved and under what conditions they are achievable, and the degree of certainty and probability with which the predicted outcomes may be observed. Ultimately, it is hoped that the mathematical formulation used to study quantum systems can address in inherent difficulty that exists today in addressing demand response where both phase (in the time domain) and duty-cycle (in the frequency domain) must be considered simultaneously. While it has not yet been satisfactorily demonstrated that the market-based control systems are quantized in the sense that "energy" levels are raised or lowered in discrete steps, the ability of the wave-function to carry both phase and frequency information is very attractive, in spite of the difficulty of its treatment. There does nevertheless appear to be a discrete aspect to transaction behavior, and the challenge is to connect this with the quantum eigenvalue problem (as seen below in Figure 3.4).

A fixed demand for Q MW of power is needed for l hours. If N_0 generators of capacity Q_0 MW are competing to offer supply, the price is competitive. [a] Further suppose the marginal cost[b] of supply is $P_{\pm 0}$ \$/MWh up to their capacity of Q_0 MW, and startup costs are C_0 \$/MW. Ignoring demand limitations, at a price P the generator earns a scarcity rent[c] $P - P_{\pm 0}$ for l hours on each MW of its capacity for a total of $l(P - P_{\pm 0}) - C_0$ of profit per MW. The threshold price for any generator of capacity Q_0 to turn on is $P_0 = C_0/l + P_{\pm 0}$. If $P < P_0$, the profit is negative, and no generator would voluntarily agree to sell any

---

(a) Competitive price and quantity are those at which the buyers and sellers agree to trade. In a perfect market, it is essential to understand that when buyers and sellers cannot agree on either price or quantity, no trade occurs.

(b) Marginal cost is defined by the MIT *Dictionary of Modern Economics* (1992) as "the extra cost of producing an extra unit of output." Paul Samuelson (*Economics*, 1973, p. 451) includes "less" in the definition, which is important in our discussion. The symmetry of "cost of increased output" and "savings of decreased output" implied by the former definition fails to hold for the discontinuous, nonconvex marginal costs curves we are required to consider for power markets.

(c) Scarcity rent is revenues minus variable costs after startup costs and no-load costs are subtracted. Profits are obtained from scarcity rent.

power. On the other hand, for any price $P \geq P_0$, the generator would offer to supply only its full capacity of Q_0 MW. Therefore when $Q < Q_0$, there is no price at which supply meets demand. Of course, when $Q = Q_0$ the price is P_0. However, for $Q_0 < Q < 2Q_0$ once again there is no price at which supply meets demand. As Stoft (2002) explains,

> *The root of the problem is that production costs are not convex. Convex production costs have the property that twice as much output always costs at least twice as much to produce. [Consider when] 100 MW for two hours would cost \$50/MWh, while 200 MW for two hours would cost \$35/MWh. Twice as much is cheaper per unit, so the production cost function is not convex.*

Figure 3.4 illustrates what occurs as generators are allocated to meet a fixed demand. When all the generators of capacity Q_0 have been committed, a new price level $P_1 = C_1/l + P_{\pm 1}$ for machines of capacity Q_1 may now be allocated. This process is repeated for each Q_i by increase value of P_i, down the economic dispatch schedule until the demand is fully satisfied. Unfortunately, for perfect markets this means that most often no price exists that fully satisfies both demand and supply constraints, and the market is not "cleared." If it did clear, it would not be at equilibrium and would not satisfy the conditions for a perfect market; either some demand would not be met or some supply would be operating inefficiently.

In today's power markets, no demand may be unmet, so the markets *must* close and there must *always* be a price that clears the market. The non-convex nature of the production cost function combined with the lack of demand response inevitably leads to deviations from market equilibrium. As a result, inefficiencies are introduced, and we must choose between inefficient but reliable generation dispatch, decreased reliability, or financial risks to generators.

The question we must therefore consider is this: *if the inefficient portion of the allocation can be exploited fully by response in demand, is there a market model according to which the system also be provably more robust*? Based on the argument above the answer is positive. But before we can explore the optimal systems, we must better understand today's suboptimal systems and the strategies used to optimize them.

Given that the market must close, consider the price of power for some supply that is less than the optimum. Assuming that suppliers agree to share the part-load equally,[a] suppliers would charge the price needed to recover startup costs over the fractional quantity sold *and* the time, or

$$P_i = n \frac{C_i}{l} \frac{Q_i}{Q} + P_{\pm} .$$

(3.18)

where n is the number of generators allocated to meet the full demand Q. This means that the excess cost is the value of excess demand which we will denote V_Q, and is the difference between the cost of the part-load portion of Q above the optimal full-load condition $(n-1)Q_i$, as given by

---

(a) In reality, other operational requirements such as balancing for frequency and voltage stability and transmission security impose conditions so that neither spreading the part-load uniformly nor fully loading each generator in sequence is used. However, the costs of these services are handled separately by the ancillary services markets and can be ignored in the present situation. Therefore, distributing the part-load condition over all generators is considered the optimal strategy for revealing the value of part-load to the demand.
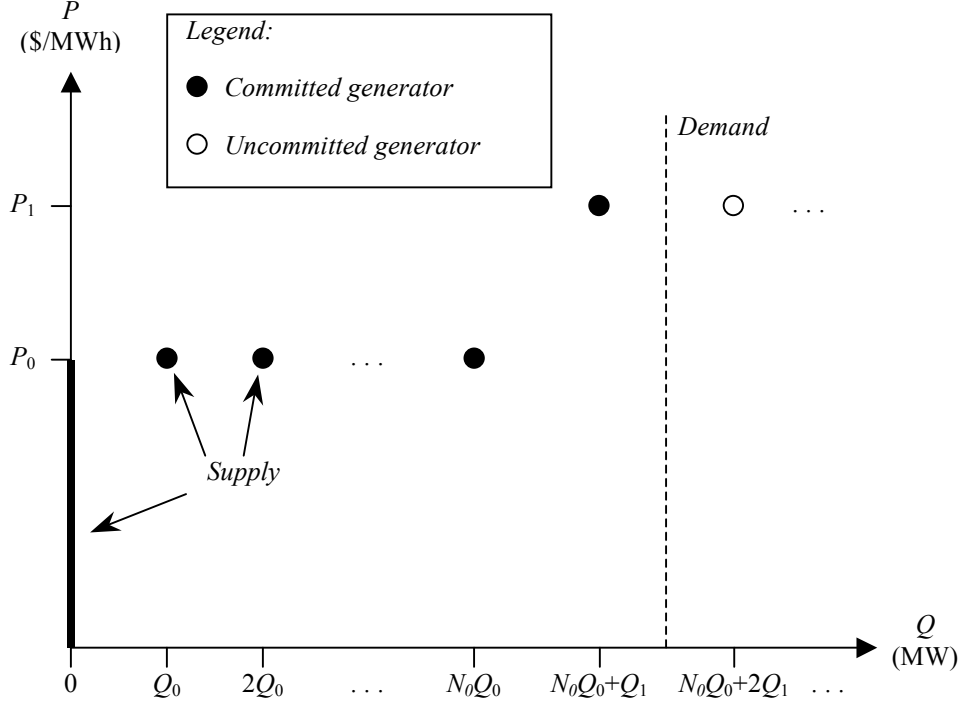
Figure 3.4. **Market State is Determinate Only for Integer Multiples of Q_i**

$$V_Q = \frac{C_i}{l}\left(n\frac{Q_i}{Q}-1\right)$$ (3.19)

This value is zero, and economic efficiency is optimal only for those demand conditions that exactly match the capacity of the first n generators of capacity C_i; e.g., $Q = n\,Q_i$ for integer values of n. The marginal value of demand reduction is the value of demand response and is found by differentiating V_Q with respect to Q, or

$$V_-(Q) = -\frac{dV_Q}{dQ} = \frac{nC_iQ_i}{lQ^2}.$$ (3.20)

For the supply curve in Figure 3.5, we see that the demand response needed to bring the system to optimal market efficiency is when $\Delta P = V_{-Q}$, the condition for which the savings is equal to the marginal value of demand response.

We observe that unless the demand response is sufficiently large to overcome the discontinuity in the cost curve at $(n-1)Q_i$, the nearest achievable optimum is at nQ_i. Therefore, one must conclude that when the magnitude of the available demand response is less than a critical threshold, $\Delta Q = Q - (n-1)Q_i$, any attempt to increase market efficiency can only be achieved by increasing demand. From the perspective of the price of power, only demand response that is at least ΔQ behaves like increased supply. Moreover, if we extend the logic, the demand response must be at least Q_i to ensure that the market will close at $(n-1)Q_i$ for all Q. Hence when the available demand response is Q_i or greater, we have a quantum market, in

Figure 3.5. Only Large Enough Demand Response Acts Like Supply Elasticity
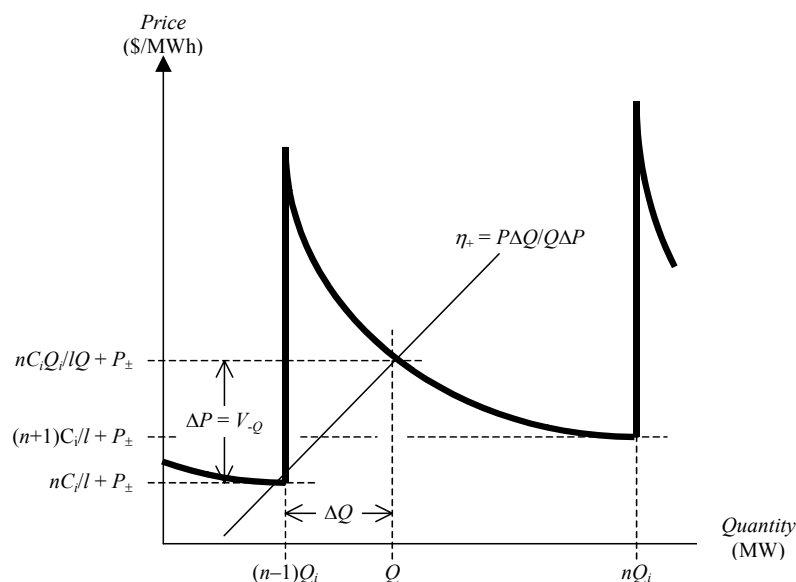
which the competitive prices and quantities are always ideal and the market is determinate. We therefore obtain the follow prerequisite for quantum market behavior:

> *The magnitude of the demand response for any load condition in a quantum market must never be less than the capacity of the last generator on the economic dispatch schedule for that load.*

The failure of any market design to meet this requirement in any part of its load regime will cause it to close inefficiently when the system demand approaches that load. Furthermore, any system whose control is based on the presumption of a quantum market will operate outside the boundaries of robustness for which it was designed. Conversely, any free and open market operating within this constraint *is* guaranteed to close efficiently.[a] This is a transactive system.

An interesting corollary is implied by this conclusion: if the magnitude of the demand response fails to exceed the capacity of the most costly generator on the current dispatch schedule, it is more economical for consumers to use more electricity if they can do so productively. The implication is that load-shifting strategies such preheating and precooling may provide better cost-reduction opportunities than energy efficiency alone, thereby reducing the magnitude of the load-curtailment needed to provide adequate demand response. However, the difficulty of the market design problem is significantly increased because searching for the lowest cost would necessarily result in increased consumption, never being able to climb over the marginal cost discontinuity at (n–1). It is thus absolutely essential to design a market that does not suffer from this shortcoming.

From a game-theoretic standpoint this may be insurmountable with any current market design. Demand elasticity, if large enough, creates an uncertainty in which strategy is optimal for any consumer. Normally, increased load will contribute to lower prices except when the increase causes another plant to

---

(a) If the market fails to close optimally, this may be taken as an early warning of market manipulation. It is this sort of behavior that should be utilized by market monitors to verify open and free market operations.

be dispatched. So, decreased consumption almost never results in a reduced price unless the decrease is large enough to cause a plant to not be dispatched. The probability of a small decrease is much greater than a large one, so without additional information or a more complex market design, consumers will *never choose to reduce* consumption.

Let us consider three common market designs to understand why they fail to produce the desired result. In the case of bilateral market equilibrium, consumers (possibly aggregated) arrange for supply at a variety of prices until the demand is met for the duration required. At some point no further trades can be completed profitably, taking into account both the consumers' value and suppliers' profits. However, because of bargaining problems, the market may never reach this equilibrium. The cost of gathering information, the incompleteness of the information available, and the time constraints imposed on the process prevent all the participants from determining which of all the possible trades are most profitable. So even if equilibrium is reached, there is no guarantee that it is efficient. Stoft (2002) argues that these problems are negligible for large systems, but does not acknowledge two factors: 1) the important role market transparency plays in the willingness to participate; and 2) the ability of small consumers to participate profitably. For our purposes though, bilateral markets use an incomplete search strategy and thus are ill-suited to the need.

In the case of exchanges, only one price is determined and is subject to auction with rules designed to limit the number of strategies and ensure that at least one Nash equilibrium exists. The rules require that prices are rounded to the nearest penny and the quantities are traded in whole MW increments. Although Stoft (2002) suggests that the rules do not appear to be required to guarantee equilibrium, it is clear that the equilibrium may not clear the market optimally. Either demand is unmet or supply is inefficient. In either case, the conditions for transactive control are not met, and thus exchanges will not satisfy the need. A power pool equilibrium is also defined as a Nash equilibrium. The main difference with exchanges is that suppliers may be provided with side payments in excess of the nominal pool price. These side payments are used to overcome that startup and no-load costs of adding new generators to the supply. However, by masking these costs from the consumers, the pool has in effect eliminated the signal that provides incentive to consumers to adjust demand. That lack of a feedback signal for the value of demand response is fatal to the design of transactive controls.

We therefore conclude that the three major market designs fail to create the conditions required to allow demand response to yield its potential benefit. We must therefore search for an alternative market design that does not prevent demand response from benefitting the market.

Another important corollary to the prerequisite for quantum markets is found when considering the demand response required when larger plants are at the bottom of the economic dispatch schedule. Clearly, more demand response is needed than when smaller plants are at the bottom of the dispatch schedule. In particular, because of their small size, distributed generation must be available in every region of the economic dispatch schedule. This results in maximum market efficiency with a minimal requirement for demand response.

Let us consider Eq. 3.21 used to generate the curve in Figure 3.5, which states that the marginal value of savings of the current demand is proportional to the inverse of the part-load demand:

$$V_Q = \frac{C_i}{l}\left(n\frac{Q_i}{Q} - 1\right).$$

<div align="right">(3.21)</div>

At any given time t, the competitive quantity of power Q is constrained to move within a certain range centered on the nominal demand. Because the demand response is limited by the design of the load controls, Q can never go beyond the bounds of this range. The problem is to determine the value of the actual demand at any time. More precisely, we would like to determine the probability of finding the competitive quantity of power between Q and $Q+dq$, at time t. When the demand response is large enough, the system is sudden jump; therefore, we express the state of the demand as a yet-unknown function $\psi(Q,t)$, which we must constrain so the system is in realizable states at all times. We express the uncertainty using Born's statistical interpretation of a wave function:

$$|\psi(Q,t)|^2\,dQ = 1,$$

<div align="right">(3.22)</div>

where ψ is not very different from a field amplitude (e.g., electric, magnetic) and thus does not force the quantum model.

This expression formalizes the indeterminacy of markets discussed above. This indeterminacy is not readily observable in ordinary market operations, considering the price of power being sold by a generator: marginal cost, startup costs, etc. However, if that information is not available, you must *buy* some power to get it; but doing so changes the state of the generator, and what you have measured is the price of your power purchase. Indeed, we showed earlier that the price would vary according to the quantity purchased. In other words, your measurement of the price interfered with the system being measured and caused the price to become determinate. What we saw earlier as mere accident of an idealized theory of markets turns out to be more fundamental. Ordinary probability theory cannot account for this, so we turn to quantum probabilities to explain the evolution of system states and its effect on the macroscopic behavior of a system.

Now we see that the problem statement as presented is incomplete. We showed how price changes with quantity according to the supply curve. But in an indeterminate system we must know how quantity changes in response to price according to the demand curve. Three possible situations should be considered:

1. Demand quantity distributed randomly within a fixed range of response; this is how most load forecasters work.

2. Demand quantity follows the marginal cost function to the local minimum. This is how today's power markets work.

3. Demand quantity seeks to the lowest price within a fixed range of response regardless of the shape of the marginal cost function. This is how a transactive system works.

We know that transactive systems are ideal, or quantum, markets, and therefore the only determinate prices are those for which generators operate at their ideal performance point. We represent this price using algebraic formulation that has a discontinuity at every determinate state of the system. As a

practical matter, this representation of the state of the system is continuous. This can be addressed by restating the price of power:

$$P(Q,t) = C_0 \left(1 - \frac{Q(t)}{\sum_N Q_{n[Q(t)]}} \right)^{-\frac{1}{\eta}} + i \left(\frac{n[Q(t)]C_{n[Q(t)]}Q_{n[Q(t)]}}{lQ(t)^2} \right) \tag{3.23}$$

where η is the supply elasticity. Therefore, it is entirely equivalent to express the price of power as

$$P(Q,t) = |P(Q,t)|e^{i\theta[Q(t)]} \tag{3.24}$$

where

$$|P(Q,t)|\cos\theta[Q(t)] = C_0 \left(1 - \frac{Q(t)}{\sum_N Q_{n[Q(t)]}} \right)^{-\frac{1}{\eta}}. \tag{3.25}$$

The advantages of this formulation are twofold: 1) it allows an explicit expression of supply response and demand response and 2) it transforms a discontinuous into a continuous model price function. Thus, we will express the market price in the compact form

$$P(Q,t) = S(Q,t) - iD(Q,t) \tag{3.26}$$

where S is the supply response and D is the demand response. As long as D is large compared with the magnitude changes in S for all Q, P is continuous for all Q.

In Figure 3.6, as $Q(t)$ increases the angle θ slowly increases, while $|P|$ remains constant. Eventually, the demand response is insufficient at price P to continue displacing supply, and the system suddenly shifts to the complementary position at angle $\theta - \pi$ but with the same price, P. The only difference is that one additional plant is now on line, and the demand response that was deferred is now made whole.

From a market-based control standpoint, this simplifies the problem of guaranteeing clearing the market. As long as the availability of demand response is present, the market will be more strongly attracted to the $n-1$ allocation of generation. However, as the availability of demand response drops, the price will go up until it is sufficient to bring the nth generator on line, when the price jumps slightly and the demand response is made whole, as shown in Figure 3.7.
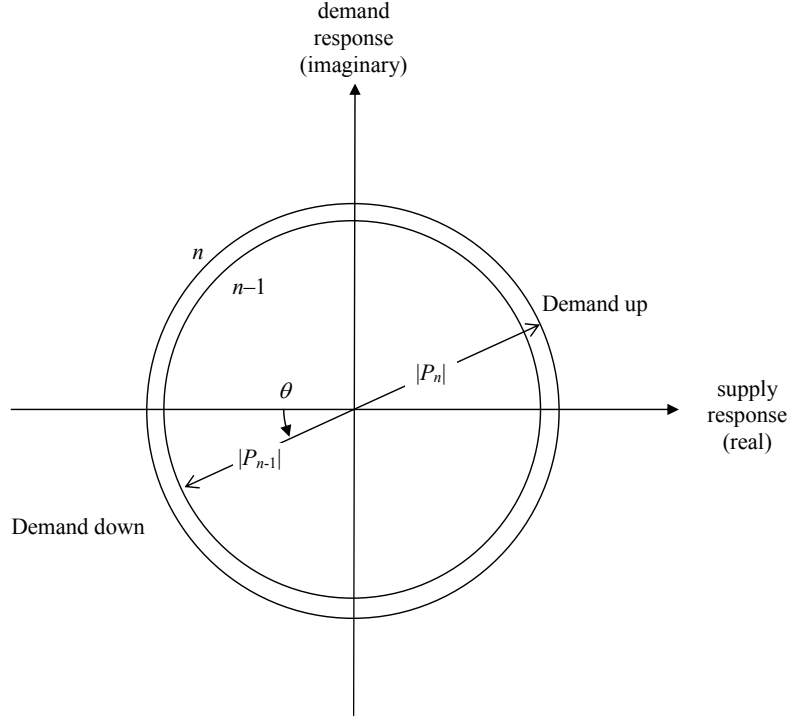
Figure 3.6.   Market Response

3.1.6   Formal Models of Load Behavior

The dynamics of electric devices and their supporting power networks are perhaps more accurately modeled using a quantum paradigm than classical physics such as those based on thermodynamics, Ising models, or even agent–based simulations. A major hurdle is defining the appropriate parameter space over which a potential or probability function can be integrated. We show how the state space of a pool of identical water heaters can be modeled as a torus. This basic model is extended to devices with different lengths of duty cycle by adding a thickness to the minor section of the torus. Demand imparts a continuous transformation of the state space in both models.

3.1.6.1   Motivation for Formalism

The probability that a lone electron moving freely along a straight line will be observed at time t_0 between two positions α and β, with $\alpha < \beta$, is typically expressed in terms of an integral whose integrand ϕ is a complex-valued wave function:

$$\int_\alpha^\beta \phi(x,t)\phi^*(x,t)dx \tag{3.27}$$

where ϕ^* denotes the complex conjugate of ϕ. Once such a representation is agreed upon, it can be extended to two or more dimensions and more complex geometries, and refined by using a suitable wave equation to select only relevant choices of wave functions.
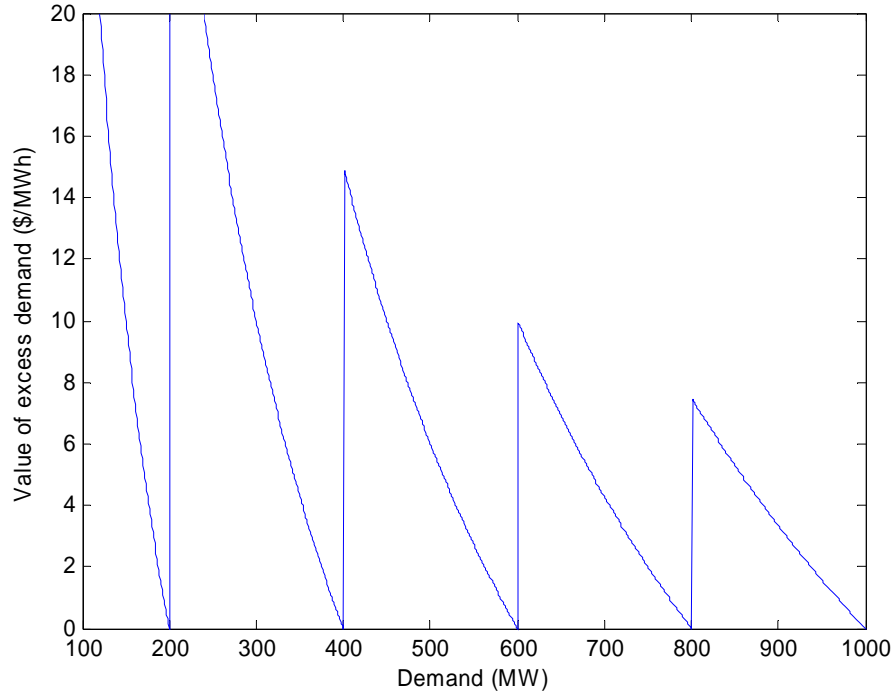
Figure 3.7.   Value of Excess Demand

The relevant questions to ask appear to be, What is *x*? What is an appropriate domain of integration? What are the appropriate wave equations? Once the questions of what is observable are answered, we can determine what forms of potentials, or pricing structure, are controlling parameters of the whole system.

The object of our model in this study is water heaters. Each device draws electricity to keep the temperature of the water within a set temperature range [$T_n...T_f$] called its *deadband*. For simplicity, the devices are identical and their deadband is a determined and invariant property. Classical mechanics defines a force by its effect on the object. The representation works when over short time spans the combined effect of forces are a linear combination of independent terms. Small forces produce continuous transformations of unconstrained state spaces. The most evident action on a water heater is that its reservoir is drawn upon by user demand. This demand either turns the subject device from *off-* to *on*-mode and decreases (temporarily) the heating rate of the device if it was already in *on*-mode. Although the time and volume of the demand are random events, in an ideal model the effect of the demand should be at least continuous and preferably modeled by a low degree polynomial or a well-studied function such as a sinusoidal or an exponential.

This section illustrates finding a system of coordinates within which the action of demand is continuous. Important characteristics of the resulting state space are the existence of a norm and of an associated topology that is rich enough to allow for integration and differentiation, which are essential ingredients in many deterministic optimization algorithms.

Time

Temperature — Consumption
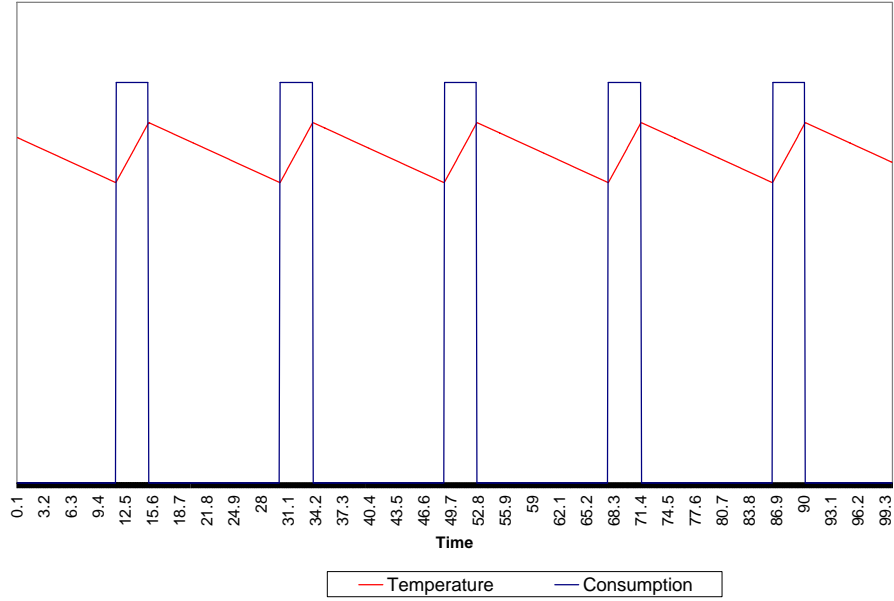
3.1.6.2 **Figure 3.8. For a Simple Device at Rest, the Consumption and Temperature are Closely Related Characteristic Waves**

At rest, that is, in the absence of demand, each device heats a fixed volume of water to a given T_f (where f stands for *off*) and then lets the water cool down back to the minimal temperature T_n (where n stands for *on*). Each heater corresponds to two periodic curves; one gives the power consumption at each instant and the other the reservoir temperature as a function of time. At rest, those two graphs closely mirror each other, as shown in Figure 3.8, which depicts a single simulated water heater. Either graph can be taken as a description of the device in the absence of any external force. The periodic square wave in the consumption graph of a device is called its (characteristic) wave.

Demand acts by perturbing the temperature and consumption cycle, as shown in Figures 3.9 and 3.10. Given that the device at rest is determined by its square wave, shown in Figure 3.8, the irregular (square) graph in Figure 3.10 corresponds to a trajectory of the device within its state space. That is, each point on the curve is a sample made at a given instant of a moving periodic square wave. There is an uncertainty built into the model because, with the basic waves being square, several motions could result in the same curve over time.

A constant positive demand makes it unlikely for a device to ever reach the lowest permissible temperature, T_n. When the demand is high enough, almost every device stores water at its maximum temperature T_f. The graph of temperature converges toward a flat line at $T = T_f$, but it never reaches it because infinitesimal temperature increments are not possible in practice. Figure 3.9 illustrates this point well: demand does not affect the slope of the upward and downward curves. Demand has two effects.
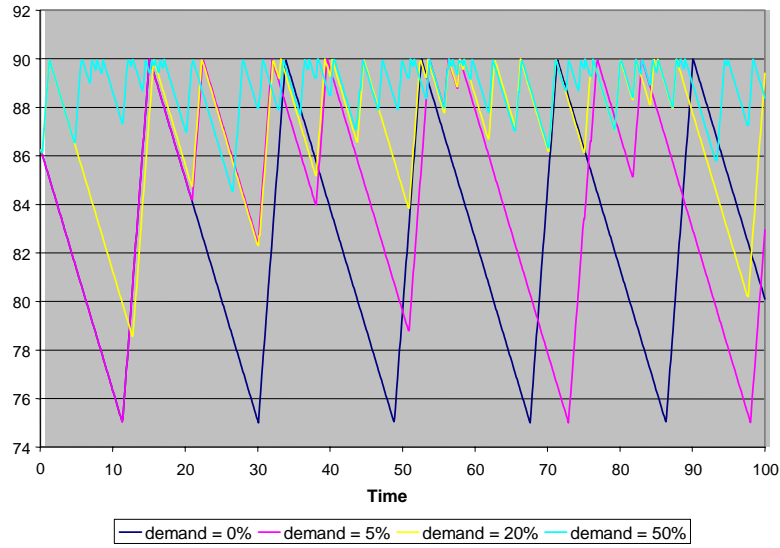
Figure 3.9.   Effect of Demand on the Heat Curve of a Simulated Water Heater

First, convergence seems to be systematically toward the same line as soon as demand is positive; second, the magnitude of demand impacts how quickly the temperature curve converges toward its limit, at least in a discrete state space.

A possible model of the effect of demand is a contraction of the duty cycle. For instance, the *x*-coordinate of a square wave might be its phase (the location of a device within its duty cycle at time zero), and its *y*-coordinate might be the length of the duty cycle. In this representation, demand draws devices toward the *x*-axis, possibly along the *y*-axis. In this visualization, demand is analogous to a pressure applied to the initial system, which springs back to steady state once the pressure is removed. This can be verified experimentally, with parameters derived, but again pressure (demand) can vary only between 0 and 1. Figure 3.11 shows the number of iterations required for a simulation to reach steady-state under zero-demand after it as reached its steady state under a positive value of demand.

Each curve corresponds to a distinct value of r_n/r_f. It is clear that irregularities occur when the cooling rate is close to the heating rate. This representation has several drawbacks. Only part of the *xy*-plane is feasible, which is counter-intuitive given how simple the system is. Another serious drawback of this approach is that, in simulations, demand decreases the effective heating rate. In fact, when demand is 100%, both the effective heating and cooling rates are zero. This is shown in Figure 3.12 using a pool of simulated water heaters.

3.17
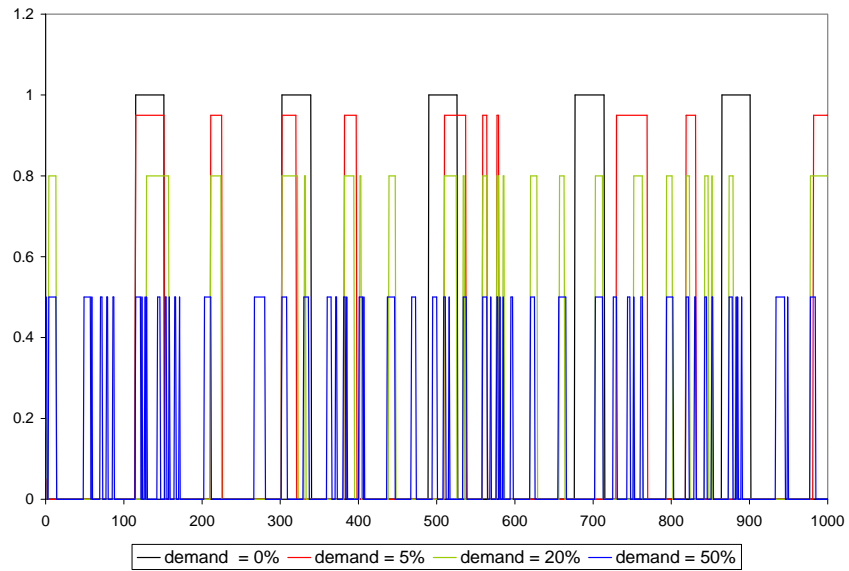
Figure 3.10. **Effect of Demand on the Consumption Curve of a Simulated Water Heater. The 0/1 curves were rescaled to improve legibility**
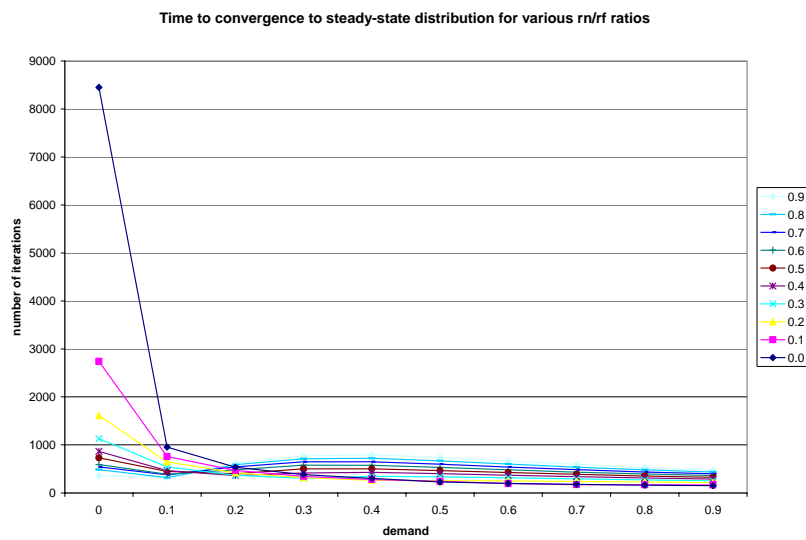
Figure 3.11. **Time for a Simulation to Reach Steady State after Demand Drops from Initial Value**
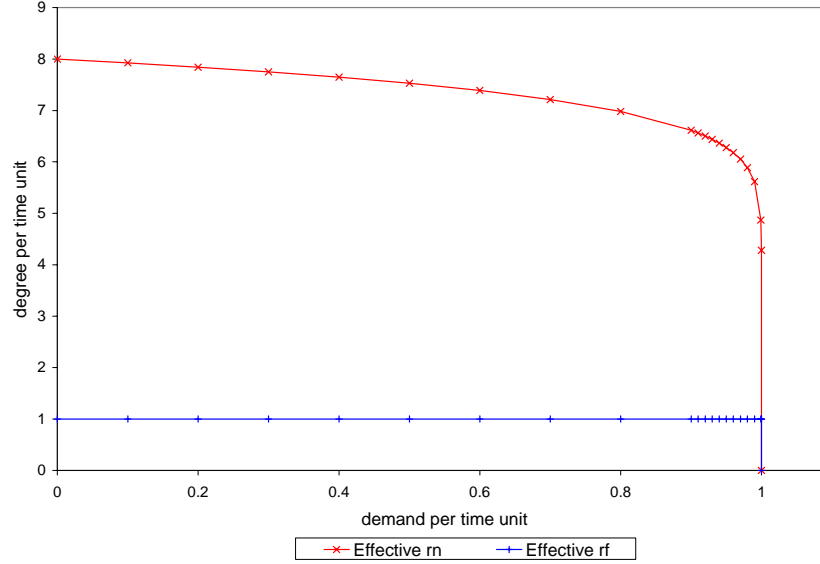
Figure 3.12. Effective Heating and Cooling Rate of a Simulated Pool Using r_n=8 and r_f=1

3.1.6.3 The $\mathbb{T}^1$ model

A device is completely determined by five numbers. The first three are hardware characteristics and are assumed identical for all devices in the pool. They are the period, *p,* of the characteristic square wave, the length, *i,* of its impulse (that is of the power demand) relative to its period, and the height, *h,* of the impulse (the instantaneous rate of power flow). The fourth quantity could be the phase but turns out to be cleaner to consider the duration, *d,* between time zero and the next impulse. The fifth quantity is the temperature, *T,* of the water reservoir, which plays the role of memory.

Our system, when unconstrained, is two-dimensional with axes *T* and *d.* Because the temperature of a device at rest increases until it reaches T_f and then decreases at another rate until it reaches the other set point T_n, a dichotomous variable can be used to record the direction the temperature is moving in. On the other hand, one can think of the temperature as the projection onto a one-dimensional axis of a point moving along a co-planar circle. The space, $\mathbb{T}$ of temperature is essentially equivalent[a] to the unit circle, $\mathbb{S}^1$; for short the $\mathbb{T}$-axis is $\mathbb{S}^1$. Figure 3.13 illustrates how the memory of a device can be mapped to the unit circle. Thus we have the relation $T = T_n + (T_f - T_n) \sin\mathbb{T}$ and $m = \text{sign}(\cos\mathbb{T})$, where + denotes *on*-mode and – denotes *off*-mode. The mode is neither *on* nor *off* when the temperature reaches T_n or T_f.[b]

Demand does not change the temperature, *T,* it changes the mode, *m,* and the delay, *d.* The effect of demand on the $\mathbb{T}$-axis is to project the *off*-mode states onto the *on*-mode states. The first step is to under-

---

(a) The exact term is *diffeomorphic;* i.e., a continuous map exists between the two spaces and it has a continuous inverse, which is all right as long as the induced topology matches our intuition about $\mathbb{T}$.

(b) In simulations with a nearly 100% demand, the temperature of water heaters reaches T_f and stays there. That is in agreement with the observed behavior of real water heaters during peak periods. The anomaly in our model at T_f is therefore not without a physical analogy. The anomaly at T_n can be justified from the symmetric case where devices would be randomly shut down.

3.19
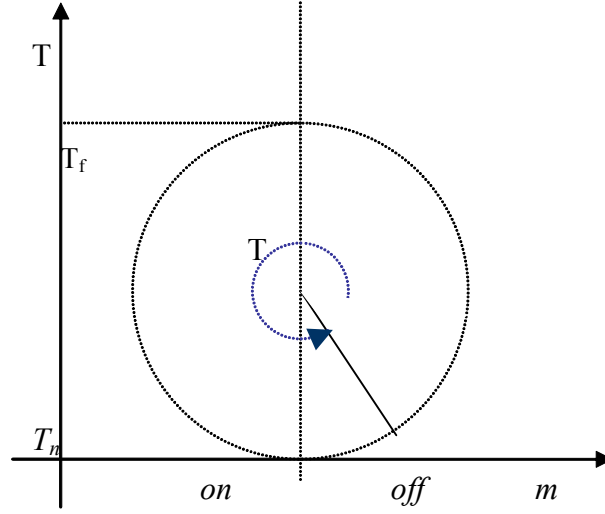
Figure 3.13. Mapping Temperature, T, and Mode, m, Coordinates into a Single Angular Coordinate T

stand how this projection affects the *d*-axis and then define a suitable topology on $\mathcal{D}$ that is a system of open neighborhoods for which this projection is continuous on the product space, $\mathcal{T} \times \mathcal{D}$.

The delay, *d,* is relative to the period of the square wave, so it is a periodic quantity although it has this unusual aspect that when it disappears to the *left* it reappears to the *right* and vice-versa. When demand hits a device in the *off*-mode, with temperature, *T*, it changes the *d*-coordinate of that device by the amount of time, Δd, needed by that device to cool down to T_n and heat back to again to *T*; the new coordinate is $d+\Delta d$ modulo *p*. The modulo operation ensures that the new value of the coordinate is between 0 and *p*. Figure 3.14 illustrates this process. If the device is in the *on* mode and the effect of demand is instantaneous, $\Delta \mathbf{d}$ is zero. IF r_n and r_f are the heating and cooling rates, both positive numbers, then $\Delta \mathbf{d}$ is defined by

$$\Delta d = \begin{cases} m = on: & 0 \\ m \neq on & (T - T_n)\left(\dfrac{1}{r_n} + \dfrac{1}{r_f}\right) \bmod p \end{cases} \tag{3.28}$$

The graph of *T* as a function of time is a curve drawn on the infinite cylinder, $\mathbb{R} \times \mathbb{S}^1$. That curve winds around the cylinder once every period. Equivalence classes of temperature curves over $\mathbb{R} \times \mathbb{S}^1$ form a torus, $\mathbb{T}^1 = \mathbb{S}^1 \times \mathbb{S}^1$, with the induced topology. It is possible to write down a formal definition of this torus in terms of quotient spaces but a geometric construction might be more informative at this stage. Take a one-period section of the infinite cylinder $\mathbb{R} \times \mathbb{S}^1$, say $[0..p] \times \mathbb{S}^1$, and glue together the two circular borders to form a torus $\mathbb{T}^1$ with the usual topology, which is not the product topology but that of a torus embedded in a Euclidean 3D-space. Any temperature curve on $\mathbb{R} \times \mathbb{S}^1$ now maps to a closed curve on the

3.20
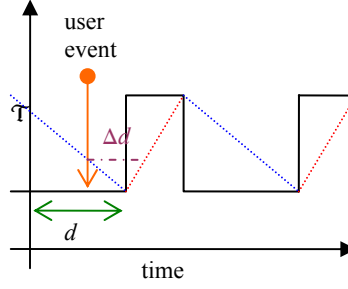
Figure 3.14.   User Demand Applied to an Off-Region of the Characteristic Wave Changes d by an Amount Δd That is Function of the Temperature T of the Device

torus that threads once through the center hole of the torus. The new representation is independent of elapsed-time.

Each point on this torus represents a characteristic curve and a point on that curve; its two angular coordinates are T along the small section and the delay, d, along the big section. A point on the torus determines the state of one device. If the inner part of the torus corresponds to the *on*-mode and the outer part to the *off*-mode, the impact of demand can be understood by looking at what happens to a small section of the torus defined by $d = d_0$. The inner semicircle of that section is unaffected; that is, the region $d = d_0$ and sign($\cos T$) ≥ 0. The outer semicircle with sign($\cos T$) <0 is stretched on the inner rim of the torus according to the process depicted in Figure 3.15, which is also along the projection of a temperature graph from the infinite cylinder onto the torus. That mapping of a small section of $\mathbb{T}^1$ can be realized continuously in 3D Euclidean space and is clearly continuous on each small section (or fiber). What makes the mapping continuous over the whole of $\mathbb{T}^1$ is that it is invariant under all rotations around the major axis of the torus.

This representation is easy to grasp visually, and a large body of knowledge already exists for dynamic systems on the torus. On the other hand, it is possible to replace the fiber $\mathbb{S}^1$ representing $\mathcal{D}$ by another more suitable space.[a] The assumption that all devices should have the same period can be relaxed to the case of bounded period devices; that is, when there is an upper bound on the period. In that case, the space of delays, $\mathcal{D}$ can be modeled not by a circle, $\mathbb{S}^1$, but by an annulus or a complete disc for devices in perpetual use. The action of demand maps inner layers of this thickened torus into themselves and is continuous. Another possible approach is to encode the length of the duty cycle as a speed.

(a) Recently, there has been some interest on modeling systems with hierarchical energy landscapes, such as spin glass formation and the binding of bio-molecules, using p-adic analysis and p-adic diffusion processes. The graphs in Figure 3.9 suggest some degree of self-similarity which is characteristic of continuous p-adic functions. The above model can be modified along those lines; rather than relying on trigonometry when modeling the delay space, one retains the modulo-arithmetic. The mapping from (T, d) to $(T, d+\Delta d)$ mod p when restricted to rational numbers (ratios of integers) induces for each T, using a standard embedding of the interval $[0, 1]$ into $\mathbb{Z}p$, a continuous transformation of $\mathbb{Z}p$ into itself. That mapping has a representation in terms of an infinite series of binomials, approximations of varying accuracy can be obtained by truncating this series.
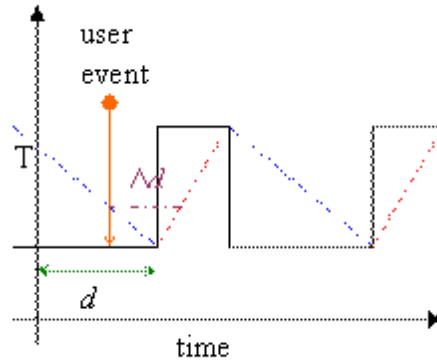
Figure 3.15.   User Demand Applied to an Off-Region of Characteristic Wave Changes, đ by an Amount, Δđ that Is a Function of Temperature, Ŧ of the Device

In the $\mathbb{T}^1$-representation, devices that are at rest (no demand and no external action) are represented by points moving along a minor section of $\mathbb{T}^1$. When r_n equals r_f and the distance traveled along the minor axis is proportional to the elapsed time. The same property can be maintained when rates differ by using an ellipse instead of a circle for the Ŧ-axis (minor) axis. With this modification, the period can be inversely related to the speed along the Ŧ-axis. Motion along the d-axis is produced by shifting the dead band of the device. Devices that are nearby in this parameter space draw power at about the same time. The purpose of control is thus to diffuse clusters of points away from costly configurations such as circles along any one of the coordinate axes.

3.2   Binary Player Model

The introduction of small-scale generation units such as fuel cells and microturbines near the point of use has been studied extensively in recent years (Edwards et al. 2000). Uncertainties linger regarding the stability and reliability impacts on transmission systems with high penetration of distributed generation (Guttromson 2002). Yasuda and Ishii (2002) proposed using Hopfield networks based on the Ising spin model in the limit case, where all the generation is provided by distributed resources, which they refer to as a "*super distributed energy system.*" They demonstrate statistical mechanics in consideration of the complex behavior of customers in such systems.

Models such as the Ising model permit consideration of the most general dynamics of highly distributed power systems in the context of bulk-power markets. However, it is difficult to interpret the results obtained from this type of model without an underlying theory that defines the meaningful aggregate properties of the system. These properties must relate (in a rigorous way) the behavior of the individual agents to the macroscopic behavior of the system as it forms its own local market or interacts with an external market. Market-based control of electric power systems has been shown to be treatable using statistical mechanics (Chassin 2004a), an easily understandable approach with an excellent record of success in physics (Kittel 1969). System entropy is a readily understood concept that only needs a brief review here. The value of trading activity in an analogy to temperature and the partition function are new to the treatment of distributed power systems, and we address them in some detail for a better understanding. Later, we discuss the consequences of considering markets as analogous to magnetic systems

and reconsider them in the context of the Ising model to uncover what general behavior may be of interest from the standpoint of control stability and robustness.

The objective of this effort is to reveal phenomena that were not immediately obvious when we first considered the behavior of highly distributed energy systems. We show how critical phase changes in market systems relate to the existence and persistence of large-scale structures. The properties and phenomena described can lead to a better understanding of this highly distributed control problem, which is neither open- nor closed-looped in nature.

3.2.1 Entropy in Distributed Power Systems

In any consideration of the aggregate properties of highly distributed power systems, it is essential that the concept of entropy be well defined and rigorously treated. In our discussions of distributed systems we work with a simple binary buy/sell model that is analogous to magnetic spin models. To simplify the model, the demand or production of any one customer among N customers is reduced to a single unit of power $\pm q$ measured in watts. For this reason it does not matter which customers consume and which generate power; only the net difference, $Q \equiv 2mq$, between the number of buyers and sellers is important, where m is the number of customers who must switch state to bring the system to a condition where $Q = 0$. The difference, Q, is the amount of power exchanged with the bulk system, and $2m$ is called the excess product of the system. The statistical weight g of any given configuration, such as that shown in Figure 3.16, is determined by the binomial distribution for any two-state system. For very large systems where $N \gg 1$, we approximate the binomial distribution using the Gaussian distribution:

$$g(N,m) \approx g(N,0)e^{-2m^2/N} \tag{3.29}$$

where $g(N,0)$ is the statistical weight of the most probable configuration of the system.

The function $g(N,0)$ is often estimated using Stirling's approximation:

$$g(N,0) \approx 2^N \sqrt{\frac{2}{\pi N}} \tag{3.30}$$

For a system where every customer is as likely to be a buyer as a seller of power, we find that the most probable configuration is the one for which $m = 0$ and the excess product is zero, i.e., there is an equal number of buyers and sellers. It may seem strange to assume an equiprobable generation versus consumption distribution, but a system with no external source of bulk power is the only configuration that is stable over time. Without an equal number of generators and consumers, the power system would not be in balance, and the frequency would deviate exponentially from the nominal frequency. This is the only configuration that clears the local market without bulk market trades, and thus the price at $m = 0$ is the clearing price.

g(8,0) = 70
σ(8,0) = log 70 ≈ 4.25
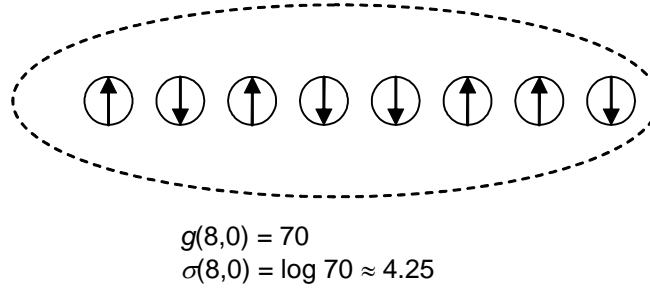
Figure 3.16.   The Most Probable Configuration of Eight Abstract Machines

In previous discussions on the subject and particularly with respect to market-based power system control, 0, entropy is always defined as the logarithm of the statistical weight of the configuration of the system. We also know that $m = Q/2q$. Therefore, we find that the entropy is

$$\sigma(N,m) = \rho(0,m) - Q^2/Nq^2 .$$
(3.31)

While the mean Q for the most probable configuration is obviously zero, the root-mean-square value of Q is given by

$$Q_{rms} = \sqrt{\langle Q^2 \rangle} = \sqrt{N}q$$
(3.32)

In other words, the fractional deviation of the system from equilibrium of a highly distributed power system increases only to $N^{-\frac{1}{2}}$ as a function of the total system capacity $Q_{max} = Nq$. We conclude that the larger the system gets, the more precisely it matches a particular expectation value of Q. It is a seemingly counter-intuitive but very important result that the most stable configuration of any isolated highly distributed system is also the one for which the entropy is maximal. This suggests that the stability of the system and entropy are closely related.

3.2.2   Value of Trading Activity

The statistical mechanical analogy can be carried one step further. The most general concept of temperature is as the inverse of the change in entropy with respect to the change in a conserved quantity, e.g., energy in thermal physics. We chose cost as the conserved value because individual agents maximize their own utility or benefit through their buy/sell decisions with respect to a relatively constant expected income. The local max-profit goal is the dual problem to the primal global min-cost problem; i.e., local profit maximization always corresponds to global cost minimization. Thus, to find a conserved quantity, we simply recognize the role the market price of power P:[a]

$$U(Q) = -QP$$
(3.33)

(a) We use the price of *power* (i.e., $/MW) and not the more commonly used price of *energy* ($/MWh). Using the price of energy requires the introduction of a time-dependence that is beyond the scope of this report. To make the two compatible in the interim we consider all behaviors as though they take place within a discrete time-step of one hour, unless otherwise noted.

$g(8,0) = 70$
$\sigma(8,0) = \log 70 \approx 4.25$
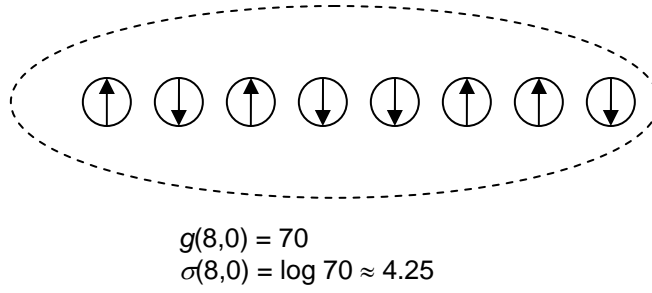
Figure 3.17. The Most Probable Configuration of 8 Abstract Machines

where $U(Q)$ is the value of the consumption (or production), i.e., the cost (or revenue).[b] (We will see later why a negative sign is introduced.) Our definition of the value τ of trading activity is based on this more general temperature concept:

$$\frac{1}{\tau} = \frac{\partial \sigma}{\partial U}\bigg|_N \qquad (3.34)$$

Any interaction between the system and an external source or sink for power is likely to change both the entropy, σ, and the excess product, Q, of the system. It is easy to see that τ is a measurement of the aggregate effect on the system of a change in the excess product of the system. During an interaction with another system, the entropies of the two systems will converge to that of the most probable configuration of the combined systems, and value will flow from the system with higher activity, τ, to the one with lower τ, regardless of system size and configuration.

The specific manner and properties of this interaction must be considered carefully to determine its impact on the systems. The behavior that interests those who study the interactions with markets, demand elasticity, is considered later. At this point we only consider the aggregate effect of market interactions as the value of τ varies, leaving the study of changes of τ in response to changes in P for a future discussion.

3.2.3 State Probabilities and the Partition Function

We continue our exploration of the analogy by adopting from statistical mechanics the notion of a partition function, Z, that provides a normalization factor for any aggregate property of the system taken over all the possible configurations of the system that are compatible with the system value U. This function permits us to answer questions such as "What is the probability of observing a local system of N customers in state l with value ε_l when connected to a market of N_0 participants with total value U_0?" We find this function by considering that the probability of a state producing or consuming the quantity Q is proportional to the total number of configurations of the system:

(b) It is not intuitive that the value U should be a conserved quantity in a system. At least over the short term, systems use the same amount of energy regardless of power price, but the power use profile itself is strongly related to the price. A higher price tends to reduce power demand briefly; inevitably though, a higher consumption follows because the total energy consumed must be maintained. This rebound phenomenon has been observed in a number of utility load shedding programs and is explained in detail by Lu and Chassin (2004).

$$\text{Prob}\{Q\} \propto g(N_0 - N, U_0 - \varepsilon) \qquad (3.35)$$

In its simplest form, we define the partition function[a] by considering that N does not change and

$$Z(\tau) = \sum_l e^{-\varepsilon_l/\tau} \qquad (3.36)$$

Any aggregate property can be estimated by computing its average over all the states that satisfy the configuration in question. For example, the per-agent average value U of the system is given by

$$
\begin{aligned}
U \equiv \langle \varepsilon \rangle &= \frac{\sum_l \varepsilon_l e^{-\varepsilon_l/\tau}}{Z(\tau)} \\
&= \frac{\tau^2}{Z(\tau)} \frac{\partial}{\partial \tau} Z(\tau) = \tau^2 \frac{\partial}{\partial \tau} \log Z(\tau)
\end{aligned}
\qquad (3.37)
$$

As an example of how this might be used, we define the *specific value C_V*, or the value a highly distributed system can absorb per unit increase of activity when connected to another system of higher τ.

$$C_V \equiv \left. \frac{\partial U}{\partial \tau} \right|_N \qquad (3.38)$$

Because we know that $U = N\langle \varepsilon \rangle$, we find that

$$C_V = N\left(\frac{\varepsilon}{\tau}\right)^2 \frac{e^{\varepsilon/\tau}}{\left(e^{\varepsilon/\tau} + 1\right)^2} \qquad (3.39)$$

When $\tau \gg \varepsilon$, the value of τ is very large compared with the size of discrete value of ε and

$$C_V \approx \frac{1}{4} N\left(\frac{\varepsilon}{\tau}\right)^2 \qquad (3.40)$$

In other words, for high levels of market activity, the value absorbed by a highly distributed energy system is inversely proportional to the square of trading activity in its local market—the more active the system, the less value it absorbs. However, the inverse is true at extremely low levels of activity, as shown in Figure 3.18. The less active the system, the less value it absorbs interacting with other systems. The value barrier in the neighborhood of $\tau \to 0.42$ is where we expect the maximum value capacity.

3.2.4   Market Interactions

In this section we consider a system of N customers, each with distributed generation capabilities as described above, with the additional proviso that they are permitted to purchase or sell any excess product

---

(a) For a more complete discussion on general partition functions, refer to Kittel,pp. 81-89 1969),.
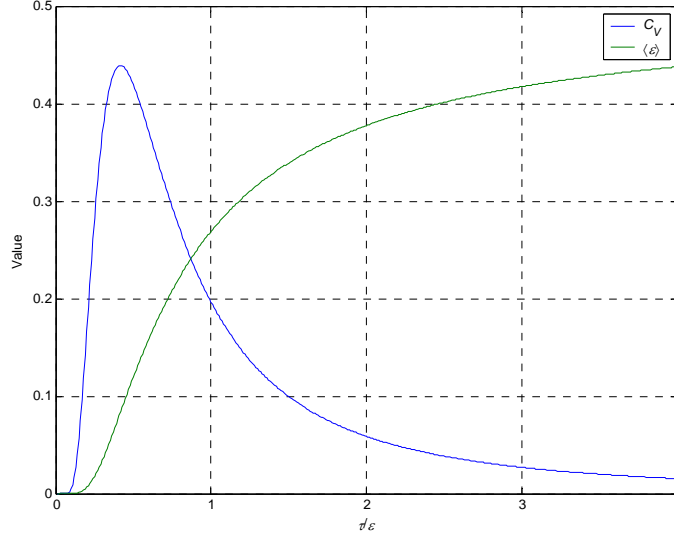
Figure 3.18. Value and Capacity of Two-State Distributed Energy Market System as a Function of Market Activity in Units of Value Separation ε

at the price P in a single market that completely surrounds the system. (We see later that the behavior of this market is highly analogous to that of a magnetic field.) The value of the interaction between a single customer and the market, ignoring interactions with other customers, is simply

$$\varepsilon_i = -q_i P \qquad (3.41)$$

where the negative sign is introduced because a system that sells a positive q decreases the net excess product it has by q, releasing an otherwise hidden potential value for the positive q. The value of all N customers is thus

$$U = \sum_{i=1}^{N} \varepsilon_i = -P \sum_{i=1}^{N} q_i = -PQ = -2mqP . \qquad (3.42)$$

According to this, the spectrum of allowed values of U is discrete, where the difference ΔU between any two values is given by

$$\Delta U = U(m) - U(m+1) = \pm 2qP = \mp 2\varepsilon . \qquad (3.43)$$

This observation should be regarded with some concern because it is reasonable to assume that such a restriction on the spectrum of values may well affect the manner in which small systems can be expected to behave when constrained.

3.2.5   The Free Value Function and Market Susceptibility

For our model, there are two sources of value that affect system entropy. We identified the internal value of the system as well as the value that results from the interaction with an external market. The interaction of these two values is something we express in terms of a single quantity, which we observe is conceptually analogous to free energy in thermal physics. The free value tells us how much value can be

transferred out of a system by interacting with it. Clearly a bulk price matching the internal equilibrium price that clears the local market reduces the total product exchange and is thus associated with higher entropy, while an extreme price (either high or low relative to the internal equilibrium price) tends to polarize the buy/sell decisions, a lower entropy state. So there would seem to be no way to extract all of the value from the system (which is perhaps reassuring). Additionally, by using the analog to free energy, we expect the free value has the following useful properties:

1. It is a minimum in equilibrium

2. It can be directly obtained from the partition function: $F = -\tau \log Z$

3. The entropy can be directly calculated from the free value.

To determine the free value (the value that can be extracted from the system by interacting with it) for our highly distributed energy system, we adopt an analog of the Landau free energy function, regardless of whether it is at equilibrium, by using a generalized entropy function $\sigma(m;\tau,P) = \tau\sigma(m)$ and defining:

$$\tilde{F}(m;\tau,P) \equiv U(m;\tau,P) - \sigma(m;\tau,P) \qquad (3.44)$$

From this definition we determine the average value of m for which the minimum condition for equilibrium is satisfied. The free value as a function of m is

$$\tilde{F}(m;\tau,P) = U(m,P) - \tau\sigma(m)$$
$$= -2mqP - N\tau \log 2 + \frac{2\tau m^2}{N} \qquad (3.45)$$

so that at equilibrium we have

$$\left.\frac{\partial \tilde{F}}{\partial m}\right|_{\tau,P} = -2qP + \frac{4\tau m}{N} = 0 \qquad (3.46)$$

or

$$\langle m \rangle = \frac{NqP}{2\tau}. \qquad (3.47)$$

Thus, in the our system we find that the free value of the system in equilibrium is

$$F(\tau,P) = -N\tau \log 2 - \frac{Nq^2P^2}{2\tau} \qquad (3.48)$$

We can verify this result easily enough if we square Eq. 3.42 to obtain

3.28

$$m^2 = \frac{U^2}{4q^2 P^2} \qquad (3.49)$$

so that when $|m| \ll N$, the logarithm of Eq. 3.29 becomes

$$\sigma(N,U) = \sigma(N,0) - \frac{U^2}{2Nq^2 P^2} . \qquad (3.50)$$

Using to the original definition of τ, we therefore find

$$\frac{1}{\tau} = \left. \frac{\partial \sigma}{\partial U} \right|_N = -\frac{U}{Nq^2 P^2} . \qquad (3.51)$$

From this we obtain

$$U(\tau, P) = -\frac{Nq^2 P^2}{\tau} \quad ; \quad \sigma(\tau, P) = \sigma(0) - \frac{Nq^2 P^2}{2\tau^2} \qquad (3.52)$$

and we find

$$Q(\tau, P) = \frac{Nq^2 P}{\tau} . \qquad (3.53)$$

We use these results, which are consistent with Eq. 3.27 to define the notion of market susceptibility $\chi \equiv dQ/dP$, or

$$\chi(\tau) = \frac{Nq^2}{\tau} \qquad (3.54)$$

a quantity that is relatively easy to observe in most power markets.

Once again, we remind ourselves that this result is only accurate for $qP \ll \tau$, because at the outset we required that $|m| \ll N$ to obtain Eq. 3.30. It would require a significantly more thorough analysis to derive expressions for both free value and market susceptibility that are accurate for cases where the market is relatively inactive and the average value of τ trading activity is small and approaches the value of a single trade qP. Nonetheless, these results are instructive in developing a more complete understanding of the behavior of highly distributed energy systems and their market interactions.

4.0   Economic Properties

Some may find it instructive to consider the relationship of market susceptibility to demand elasticity, which is defined as (Henderson and Quandt 1980)

$$\eta_D = \frac{d\log Q}{d\log P} = \frac{dQ}{dP}\frac{P}{Q}.$$

(4.1)

The demand elasticity indicates how the quantity purchased changes in relation to the price. When $|\eta_D| < 1$, demand is inelastic and the percentage change in price is greater than that of quantity. The increase in the total value U of trading volume with the external market from higher prices dominates the decrease in external trading value from the larger quantity, and the value increases when price increases. Conversely, when $|\eta_D| > 1$, demand is elastic and the percentage change in quantity is greater than that of price. The increase in trading from higher prices is dominated by the decrease in trading from smaller quantities, and the market trading decreases when the price increases. In the middle is that condition when $|\eta_D| = 1$, and the demand elasticity is unitary. In this case, the percentage change in price is inversely proportional to the percentage change in quantity. As a result, the market trading value is unaffected by price fluctuations.

We stipulated at the outset that our system was subject to three important conditions: 1) it was closed, causing the total system value to be unchanged; 2) it was large, causing N to be very large; and 3) it is near equilibrium, causing deviations in m to be very small relative to N. We find that for these conditions the demand elasticity is unitary:

$$dU = -QdP - PdQ = 0$$

(4.2)

or

$$\eta = \frac{dQ}{dP}\frac{P}{Q} = -1$$

(4.3)

Unitary elasticity is a necessary consequence of the assumptions we make, in particular, the imposition of closure. The objection is often raised that "real" systems are not closed. We assert that what is generally meant by *not closed* is simply that not all external interactions are accounted for, especially those relating to entropy. If we permit a system to interact with more than one market, the above discussion cannot explain the nonunitary elasticity of the system with respect to any one of the markets. When the elasticity of a system changes with respect to one market (e.g., for one fuel source) it must change for the others (all alternative fuel sources) in a way that preserves their combined unitary elasticity. The net result is that when all external interactions are accounted for in large highly distributed energy systems, the systems can be considered closed, and the aggregate demand elasticity over all K markets combined is always unitary when the system is at equilibrium, but there are many ways to satisfy the relation:

$$\sum_{i=1}^{K} dU_i = -\sum_{i=1}^{K} Q_i dP_i - \sum_{i=1}^{K} P_i dQ_i = 0,$$

(4.4)

4.1
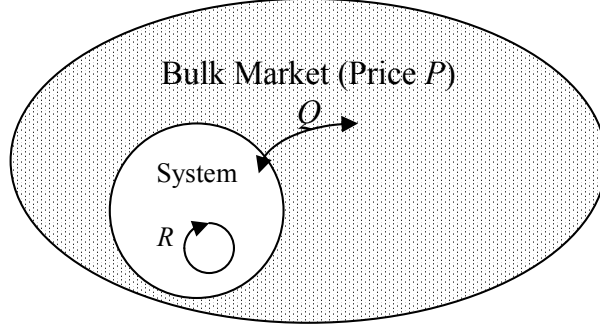
Figure 4.1.   A Local Market-Based Power System Interacting with a Bulk Power Market

while maintaining the equilibrium condition of (4.1). More interesting is the fact that given an observed elasticity η_k with respect to the kth local market, we can compute the net dU_k lost or gained to it. Given the result of (4.2) we find that

$$dU_k = (\eta_k + 1)Q_k dP_k \,. \tag{4.5}$$

4.1   Local Interactions

We denote the quantity of commodity exchanged between a system and a market by Q, as shown in Figure 4.1. Because of the small relative size of the local power system compared with the bulk market, the price P of this exchange is determined primarily by external influences and largely independent of Q. The quantity R denotes all possible mutual trading relationships in the local system. We saw that Q has a strong analog to the magnetization in a binary spin model in thermodynamics and P is analogous to the magnetic field. By analogy we propose that the Hamiltonian for such systems is given by

$$U = -QP + \frac{1}{2}BR \tag{4.6}$$

where B is the local exchange price of power. It was previously given that $Q = 2mq$ and we show in Figure 4.2 that $R = (4m^2 - N)q$. The vernacular interpretation of the value $2m$ is that it is the difference between the number of buyers and the number of sellers. It is also useful to recall that m is the number of participants who must change their states to bring the system to a configuration where $Q = 0$ and is the most probable configuration.

If a system is connected without constraints to a market at constant price P, then $B = P$ and substituting into Eq. 3.42 we obtain

$$U = Pq\left(2m^2 - 2m - \frac{N}{2}\right) \tag{4.7}$$

which we differentiate with respect to m to obtain

$$\frac{dU}{dm} = 2Pq(2m - 1). \tag{4.8}$$

4.2
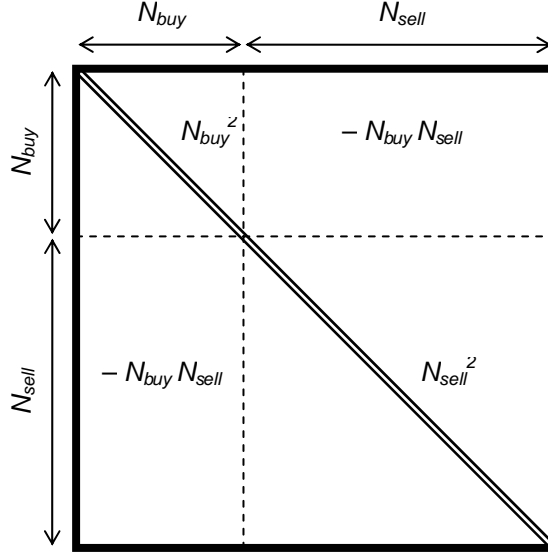
$$R^* = N_{buy}{}^2 - 2N_{buy}N_{sell} + N_{sell}{}^2$$

$$= (N_{buy} - N_{sell})^2$$

$$= (2m)^2$$

$$= (Q/q)^2$$

$$R = (R^* - N)q$$

$$= Q^2/q - Nq$$

Figure 4.2. Potential Value R^* of All Possible Local Trading Relationships in a Market-Based Power System. R excludes the value of self-supply relations

The second derivative $d^2U/dm^2 = 4Pq$ determines which condition corresponds a minimum. When $P > 0$ the curvature is positive and the Hamiltonian is minimized when $m = \frac{1}{2}$ or $Q = q$. When $P < 0$, the curvature of U is negative and the minimum occurs when $m = -\frac{1}{2}N$ or $Q = -Nq$. These results are shown in Figure 4.3. The $m = \frac{1}{2}$ solution for positive prices cannot be satisfied by any single configuration. For the local system to maintain such a minimum it must exhibit a continuous fluxing of states such that the expectation value $\langle m \rangle = \frac{1}{2}$. This leads us to conclude that the most stable solutions to local system control of small systems are ones that encourage and indeed rely upon the transactions and distributed decision-making taking place in local markets. It is also a reassuring verification that for negative prices, the most efficient configuration of the system is found when all local participants in the market are buyers and Q is at its most negative possible value $-Nq$.

4.2 Losses and Gains

The value U_Q of the exchange of Q at price P between the system and the market is given by

$$U_Q = -QP . \tag{4.9}$$

The difference between this value and the total value is the potential value, V_Q, of the local trading:

$$U = U_Q + V_Q \tag{4.10}$$

where V_Q is the potential of the local market as derived from Eq. 3.42, such that

$$V_Q = \frac{1}{2q} PQ^2 - \frac{Nq}{2} P \tag{4.11}$$

and its derivative is

$$\frac{dV_Q}{dt} = \frac{PQ}{q}\frac{dQ}{dt} + \frac{Q^2}{2q}\frac{dP}{dt} - \frac{Nq}{2}\frac{dP}{dt} \tag{4.12}$$

We call any change of total value U of the system a net gain (or loss) and denote it W. A gain or loss resulting from a change dP in price or a change dQ in quantity is given, respectively, by

$$W_P = -\int_{P_0}^{P_1} Q dP \quad ; \quad W_Q = \int_{Q_0}^{Q_1} P dQ \tag{4.13}$$

and the total gain (or loss) is the combined effects or the sum:

$$W = W_P + W_Q = -(P_1 - P_0)Q + (Q_1 - Q_0)P \tag{4.14}$$

If the changes $dP = P_1 - P_0$ and $dQ = Q_1 - Q_0$ occurs over a time interval dt, then

$$W = P\frac{dQ}{dt} - Q\frac{dP}{dt} = \frac{dU_Q}{dt} \tag{4.15}$$

The values of P and Q are observable in one-hour increments in most markets and thus we can directly compute W_P and W_Q from ISO data records, as is shown in Figure 4.4. The relative strength of the oscillatory behavior of W_Q relative to W_P and that the impulse nature of W_P is absent from W_Q. If

price and quantity were fully coupled, in responding to each other we would observe complementary behavior in the two signals. Indeed, as seen below, demand elasticity is closely related to these two quantities, and given sufficient elasticity the values of the W_P and W_Q become tightly coupled.

4.3   Demand Elasticity and System Stability

Up to now we have treated the market as a reservoir with a price P, which does not respond measurably to the quantity Q consumed or produced locally. However, in the limit, we allow the local system to grow until it becomes the entire market. Now the price fluctuations are determined by how the local quantity fluctuations combine globally. The only quantities consumed or produced are local ones, but number N of market players is very large. Previously we treated the behavior of price for a small system at equilibrium with a large market and could not consider the feedback of local price

responsiveness on the market. But now, as in the magnetic dipole model, local fluctuations of the quantity affect the market, and indeed, in the limit, determine the market. With this change in mind, we can now consider in more detail the effect on quantity changes on price.

The demand elasticity of a quantity Q is defined as

$$\eta \equiv \frac{d\ln Q}{d\ln P} = \frac{P}{Q}\frac{dQ}{dP} = \frac{P\,dQ/dt}{Q\,dP/dt} = \frac{dW_Q/dt}{dW_P/dt} \tag{4.16}$$

Elasticity is an indicator of the willingness of the purchaser to forego the transaction when the price increases. Commodities with large elasticities ($\eta < -1$) are generally considered luxuries in that an increase in price will result in a disproportionately high decrease in consumption; whereas commodities with small elasticity are considered necessities, their response being disproportionately low (Henderson and Quandt 1980). Unitary elasticity $\eta = -1$ occurs only for closed systems in which U is conserved, for which we always have

$$W_P = -W_Q \tag{4.17}$$

If the goal of creating local markets it to achieve a unitary elasticity $\eta_b = -1$ with respect to the bulk market then this condition is met when

$$\frac{dU_Q}{dt} = -Q\frac{dP}{dt} - P\frac{dQ}{dt} = 0 \tag{4.18}$$

It is easy to see that the condition for Eq. (4.17) is the same as that for (4.18). When it is not, we must provide for the response from the potential V_Q, based on a elasticity η_l with respect to the local market or

$$\frac{dV_Q}{dt} + (1+\eta_l)Q\frac{dP}{dt} = 0 . \tag{4.19}$$

The imbalance is obtained from among the local market's trading relationships R so that the total value is conserved, and thus

$$(1+\eta_l)Q\frac{dP}{dt} = \frac{dV_Q}{dt} . \tag{4.20}$$

In this way the local market appears to have unitary elasticity in the bulk market, so that $dU_Q/dt = 0$ and solving for P we obtain

$$P = \left[-\frac{1}{2}Q + (1+\eta)q + \frac{1}{2}\frac{Nq^2}{Q} \right]\frac{dP}{dQ} \tag{4.21}$$

the solution of which is the local market's price function depends on the value of $Z(\eta)$ where $Z(x) = x^2 - 2(1+\eta)x - Nq^2$. When $Z(\eta) > 1$

$$P \propto \frac{1}{Z(Q)} e^{-2\frac{(1+\eta)}{\sqrt{Z(\eta)-1}}\tan^{-1}\frac{Q-(1+\eta)}{\sqrt{Z(\eta)-1}}};$$ (4.22)
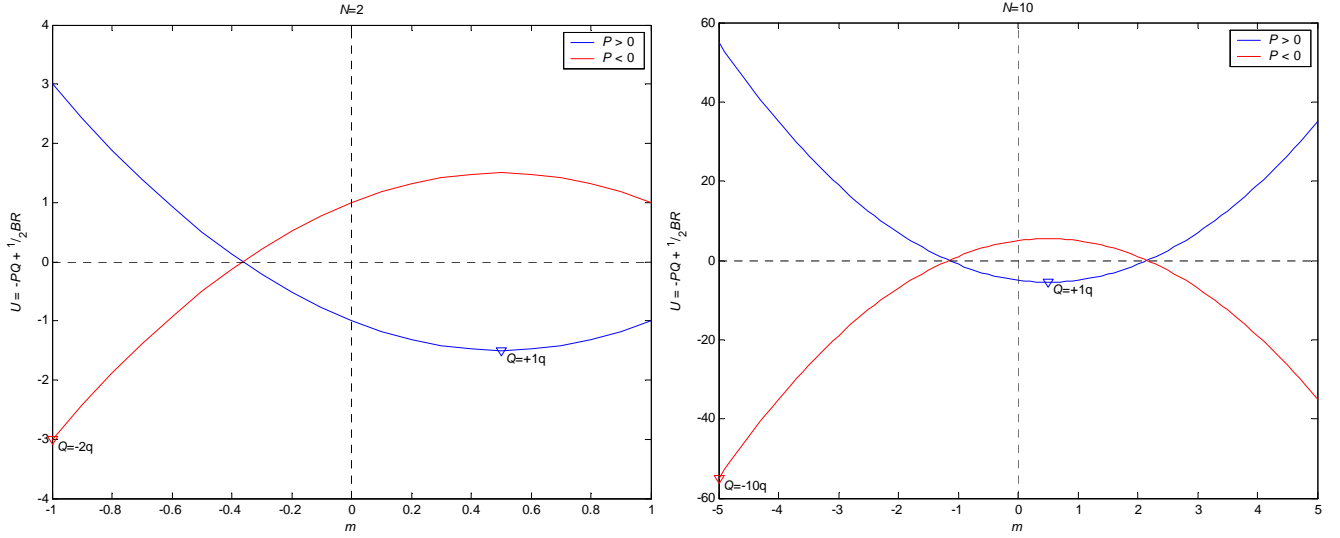
Figure 4.3. **Local Market Clearing Prices and Quantities for Systems with 2 and 10 Distributed Generating Units**

when $Z(\eta) < 1$

$$P \propto \frac{1}{Z(Q)} e^{2\frac{(1+\eta)}{\sqrt{1-Z(\eta)}}\tan^{-1}\frac{Q-(1+\eta)}{\sqrt{1-Z(\eta)}}};$$ (4.23)

and when $Z(\eta) = 1$

$$P \propto \frac{1}{Z(Q)}$$ (4.24)

This price function shown in Figure 4.5 has two asymptotes $Q = 1 + \eta \pm \sqrt{Z(Q)-1}$. In large systems, these price asymptotes are far from the theoretical limits of $Q = \pm N$. For a system with unitary elasticity the price limit occurs when $Q \rightarrow \pm\sqrt{N}$.

This result suggests that relatively small systems have a great deal of flexibility in the price for fractional quantities (Q/Nq), but as systems get larger, the price escalates for smaller fractional quantities. In effect, the price asymptote guarantees that quantity and prices will remain stable within the range $\pm Q_{rms}$. If the consumption or production must go outside that range, it will necessarily jump to $\pm Nq$. This is essentially a stability condition that is guaranteed by the presence of the local market.

4.4 The Two-Dimensional Ising Model

The Ising model was originally introduced by Wilhelm Lentz in 1920 (Lentz 1920) and the one-dimensional model of ferromagnetism was solved by Ernst Ising in his doctoral thesis in 1925 (Ising 1925; Peierls 1936). The two-dimensional problem in the absence of an external magnetic field was first solved by Kramers and Wannier (1941), Onsager (1944, Kaufman and Onsager 1953), and later refined by Feynman and others giving the same solution using more elegant and approachable methods (Feynman 1972). Many thousands of papers have been published since then on the subject, making it one of the most studied problems in statistical mechanics. One particularly interesting extension to the Ising model was studied by Potts (Wu 1982). The Potts model allows for more than two states. While the closed-form solutions of the original Ising model are useful from a pedagogical standpoint, their relevance to more realistic problems may be limited; and the dearth of approachable closed-form solutions for both Ising model in non-zero magnetic fields and the Potts models in general have been perceived as an obstacle to their general use (Istrail 2000; Dorogovtsev et al. 2003).

According to the Ising model, the interaction of each element (in our case a market player) is described by two distinct terms: the interaction with the external market and the interaction with its immediate neighbors. By disaggregating Eq. 3.42 we easily see that the total value, ε, of the market interaction of a single customer, I, is $-Pq_i$. However, under the Ising model, only the interactions of customer i with her immediate neighbors are considered, giving us the system's formal Hamiltonian:

$$U = -P\sum_{i=1}^{N} q_i - \frac{J}{2}\sum_{j=}^{N} q_i q_j \; .$$

(4.25)

The value J determines the value of the coupling between the neighbors. In the case of a positive coupling, where similar states are favored, positive values of J are used, and in the case where opposite states are favored, negative values of J are used. The sign of J describes the tendency for the system to favor longer- over shorter-distance interactions, and the value of J with respect to P as describing the relative strength of that coupling with respect to the effect of the market.

We begin by examining the two-dimensional Ising model used by Yasuda and Ishii (2002) for the zero-market Hamiltonian (Eq. 4.25). With $P=0$, the most probable configuration of the system satisfies the zero excess product at $Q=0$. Such a system behaves in two very distinct ways depending on the value of τ. The question of critical value, τ_c, at which the phase change occurs, is of particular interest to those who study Ising models. The solution has long been known in closed form for the simplest 1- and 2D systems and is called the Curie temperature in statistical physics. The existence of the Curie point requires using the free value derived above. We rely on the observations of Yasuda and Ishii (2002) to infer the necessary existence of the critical Curie point, determined according to mean field equation:

$$\langle m \rangle = \frac{\frac{1}{2}NqP}{\tau - \frac{1}{2}\alpha N}$$

(4.26)

where α originates in the mean field approximation[a] according to which the interaction value depends only on the magnitude of the excess product, $|2m|$.

The existence of the $-\frac{1}{2}\alpha N$ in the denominator of Eq. (4.26) suggests an important transition at that value of τ. Specifically, the theory of ferromagnetism predicts that it is possible for the value of $\langle m \rangle$ to be non-zero, even when $P = 0$. We conclude that while it is probable a system will have a non-zero excess product when market activity is low, once market activity rises above the critical level, it is more likely that the excess product will be tightly coupled to the price. Clearly, to achieve demand elasticity, we must not only introduce distributed generation but also ensure there is a sufficiently active market to couple the system's excess product to the price. Lower the value of trading activity and the excess product will decouple from the price; and the system will eventually collapse into a single-state configuration of minimum entropy and thus minimum value. Below the critical point, when $J > 0$, the outcome of the system will eventually coalesce into a single state, either all buyers or all sellers. However, all else being equal, it is *a priori* impossible to tell which outcome will emerge.[b] It is only by sustaining a value of trading activity above the critical point that one observes a sufficient diversity of states to considered the behavior of the system truly a highly distributed energy system.

4.5 Equation of State for Elementary Markets

The equation of state establishes the relationship between the value of trading activity and the total value of the system. To determine the equation of state we must begin by integrating Eq. (???):

$$U = \frac{1}{2}\varepsilon^2 n_T \tag{4.27}$$

We note from the definition of entropy that we have

$$\sigma = \sigma_0 - \frac{2n_t^2}{N} \tag{4.28}$$

into which we substitute the result from above to obtain

$$\sigma = \sigma_0 - \frac{1}{2}\frac{U^2}{N\varepsilon^2} \tag{4.29}$$

and therefore

$$\frac{\partial \sigma}{\partial U} = \frac{U}{N\varepsilon^2} = \frac{1}{\tau} \tag{4.30}$$

(a) The mean field approximation is consistent with the Ising model—it ignores the range dependence of the actual interaction for which the actual strength diminishes with distance. Formally, the mean field approximation requires adding the term $-\frac{1}{2}\alpha m^2$ to the free energy function, whose derivative is zero at equilibrium, leading us to Eq. (4.26).

(b) In reality, all things are never strictly equal, and it should be easy to predict which condition will emerge.

or

$$U = \frac{N\varepsilon^2}{\tau}. \qquad (4.31)$$

This result is quite different from the result expected from Sergeev's discussion of ideal systems (Sergeev 2003), which is expected because Sergeev assumed a market model analogous to the ideal gas, while this model is much more like a magnetic spin model.

4.6 Carnot Cycle

When, by means of some control strategy a small system is repeatedly brought into alternating contact with two very large systems with differing values of market activity, τ, we observe a process that results in the production of profit value, W. The effect of this process is shown in Figure 4.6.

The value W is non-zero whenever the $\Delta\tau$ is non-zero because the value absorbed by the system in the higher market is given by

$$V_{high} = \tau_{high}\left(\sigma_{high} - \sigma_{low}\right) \qquad (4.32)$$

while the value released into the system of low activity is given by

$$V_{low} = \tau_{low}\left(\sigma_{low} - \sigma_{high}\right) \qquad (4.33)$$

So, ideally, the value extracted W is given by

$$|W| = V_{high}\left(1 - \frac{\tau_{low}}{\tau_{high}}\right)$$

However, in practice it is difficult to bring the entropy of the system into perfect equilibrium with either the low or the high system without some dissipation of entropy. There it is reasonable to expect the V_{high} will be smaller than theoretically possible; therefore, W is not always as large as desired.

The concept of a Carnot cycle can be applied to this system by observing that the changes in price and quantity are related to changes in entropy and market activity, as shown in Figure 4.7.

4.7 Application to Market Monitoring

The existence of a Carnot cycle for econophysical market systems implies the existence of a maximum efficiency $V_{low}/W = \tau_{low}/\Delta\tau$, analogous to the Carnot efficiency. We conclude that the maximum profit possible is W when the losses are zero. When we observe profits in excess of W, conservation laws have been broken and one of the market "rules" have been violated, exposing the existence of a "cheater."
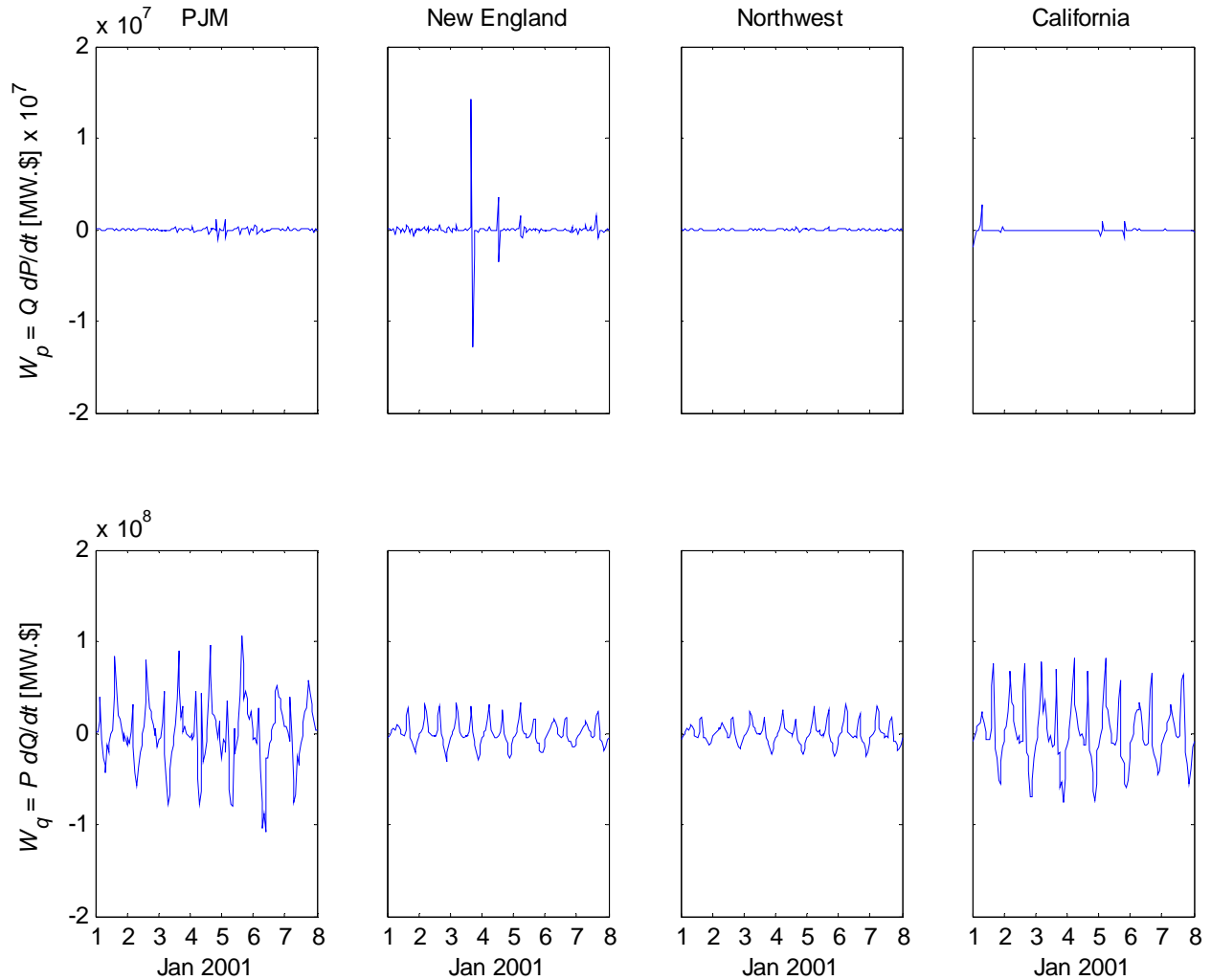
Figure 4.4. **Values of W_P and W_Q of Four U.S. Power Markets in the First Week of January 2001**

The advantage of this method is that any violation of market rules that does not increase profits above W is not a violation of the conservation laws. Without fruit it is technically not a violation at all and can be ignored. Thus it is possible to distinguish between cupidity and stupidity. The latter is most often punished by the market itself, and the former is the only one requiring follow-up by regulators.

Moreover, any flawed rule that permits a breach of a conservation law would also be easily detected in spite of the flaw. This has the advantage of detecting market problems at their most fundamental level rather than relying on intuition and extensive testing to identify the rules needed to prevent cheating and search for missing, useless, or counterproductive rules.
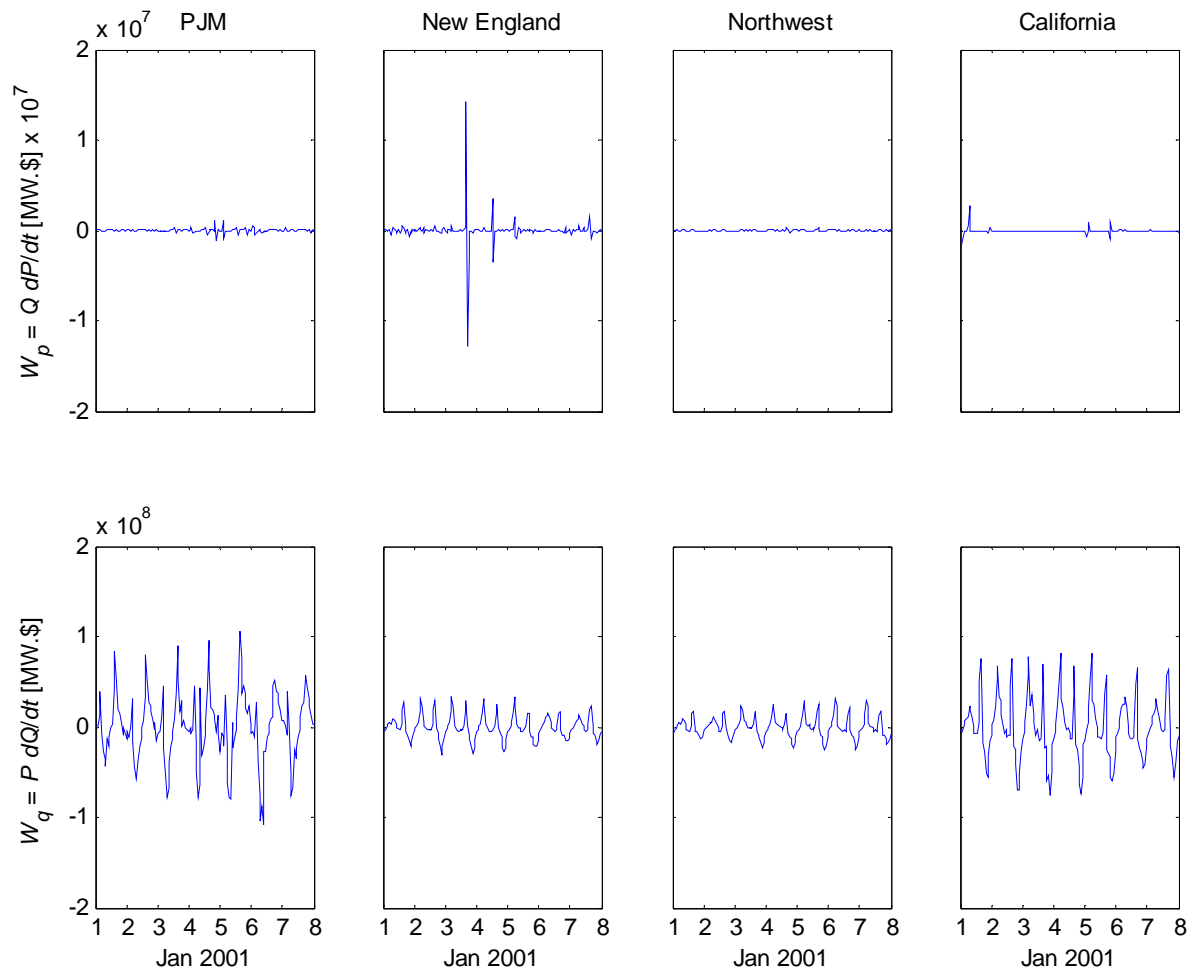
Figure 4.5. The Values of W_P and W_Q of Four U.S. Power Markets in the First Week of January 2001
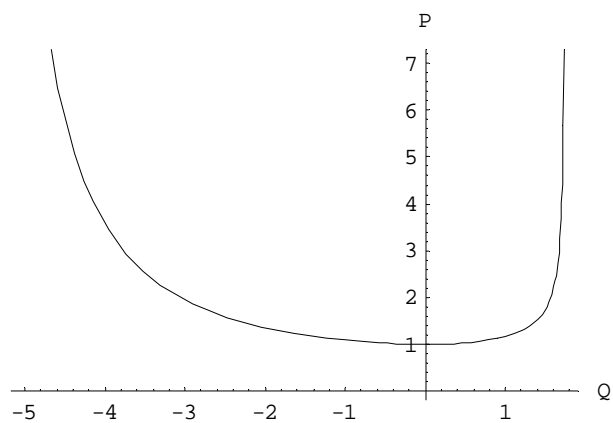
Figure 4.6. The Price Function for an Illustrative System of $N = 10$ Agents Having Elasticity $\eta = -3$
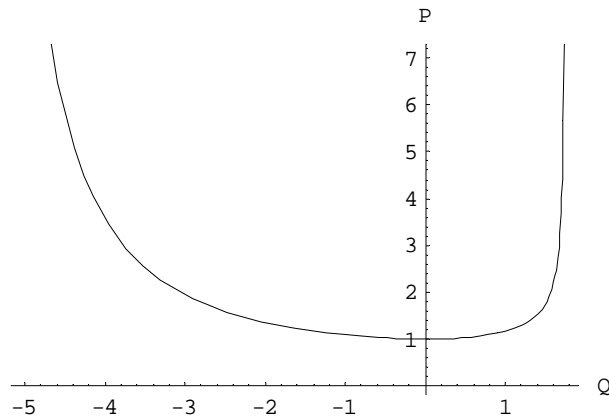
Figure 4.7. Price Function for Illustrative System of $N = 10$ Agents with Elasticity $\eta = -3$
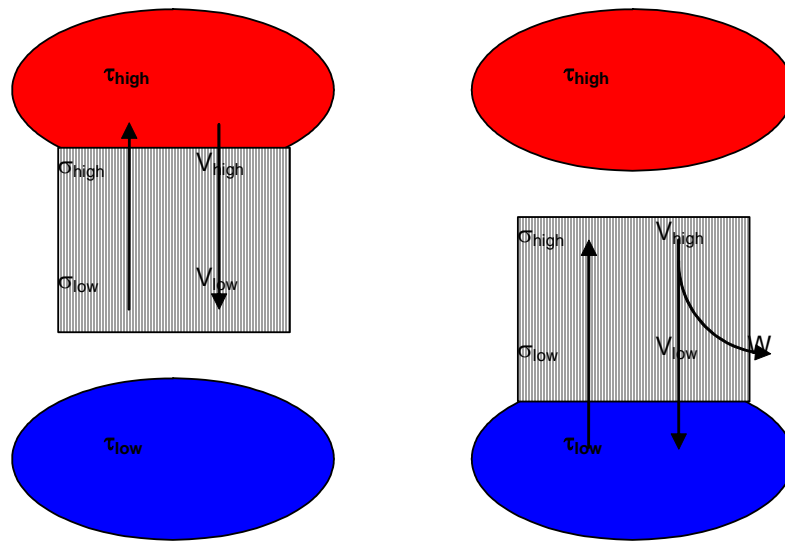
Figure 4.8. Production of Value W by a System Mediating Between Two Other Systems
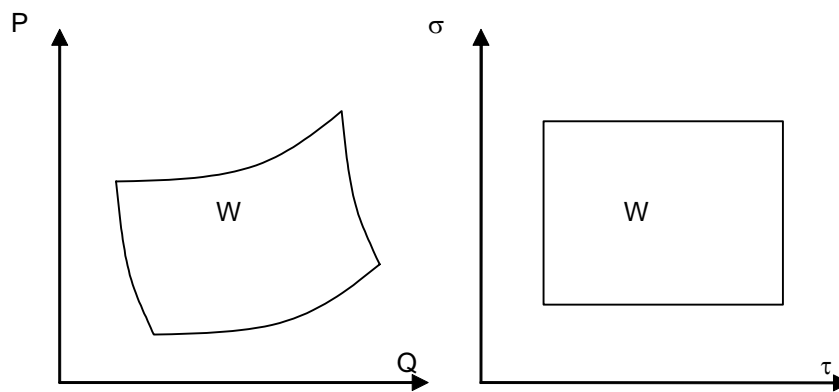
Figure 4.9. Carnot Cycle in PQ Space and in ST Space

4.12

5.0   Application to Load Forecasting

Advanced market-based load control provides numerous economic benefits to utilities, electric generation and transmission operators, and their customers (Kannberg et al. 2003). Furthermore, load control has long been considered an effective strategy in mitigate price volatility and its adverse effects on electricity markets (Eto et al. 2001). However, the interactions between markets and loads are not well understood phenomena. Traditionally, three major price responsive load control strategies have been discussed: curtailment (Bohn 1982), substitution (Schweppe et al. 1989), and storage (Daryanian et al. 1989). But, more recently, load shifting end uses that cycle naturally have been considered more carefully because load shifting is less likely to adversely affect the customer, particularly when done over relatively short time frames (Lu and Chassin 2004), and it is more readily implemented as a true distributed control strategy that requires no centralized dispatch (Chassin 2004a). The key advantage of load shifting is that it contributes positively to overall load leveling with system-wide economic benefits (Kannberg et al. 2003).

The Power Distribution System Simulator (PDSS) was developed at Pacific Northwest National Laboratory to study the benefits and challenges of this type of advanced, market-based load control strategies (Guttromson et al. 2003). PDSS is a state-of-the-art high-performance agent-based simulation capable of individually modeling many thousands of end-use loads with a wide variety control strategies, thermal models, and demand models. Demand behavior was developed from ELCAP load-shapes (Taylor and Pratt 1989), and the thermal models use the Equivalent Thermal Parameters (ETP) model (Taylor and Pratt 1988).

The validation of PDSS involves a large number of tests, each designed to verify the correctness of the code, the effectiveness of the algorithm, or the validity of the model. Problems detected during these tests are examined to determine the root cause, and corrections are implemented and retested to verify that the problem is no longer present. This paper examines two problems that were uncovered in PDSS during the course of testing. These problems were ultimately traced to a modeling error that suggested that the standard model for load diversity used by PDSS might be flawed.

5.1   Classical Model of State Diversity

PDSS makes the assumption that under steady-state conditions the probability density function for the temperature and mode of thermostatic end-use loads is based solely on the heating and cooling rates of the stored thermal mass (Guttromson et al. 2003). However, three parameters are known to govern the behavior of thermostatic loads: the duty cycle, the cycling period, and the end-use demand (Lu and Chassin 2004).

The duty cycle, φ, is estimated based on the Duty Cycle Model (Ryan et al. 1989), which states that

$$\varphi = \frac{Q_{avg}}{Q_{max}} \tag{5.1}$$

where Q_{avg} is the average connected load and Q_{max} is the total connected load. We can estimate the average connected load, Q_{avg}, based on the fraction of on time, t_{on}, with respect to total cycling time, $t = t_{on} + t_{off}$. So we have $Q_{avg} = Q_{max} \times t_{on} / t$, and we find that

$$\varphi = t_{on}/t \tag{5.2}$$

The duty cycle is a dimensionless number between zero and one. In cases where t_{on} and t_{off} are not known *a priori*, they are often estimated from the heating and cooling rates r_{on} and r_{off},[a] such that

$$r_{on} = \frac{T_{off} - T_{on}}{t_{on}} \quad , \quad r_{off} = \frac{T_{on} - T_{off}}{t_{off}} \tag{5.3}$$

where T_{on} and T_{off} are the thermostat's true on and off temperatures, which includes the deadband around the set-point.

The demand, η, is given as an exogenous parameter that describes the rate at which appliances are subjected to demand events that cause the controller to immediately transition to an *on* state. Note that η affects all appliances whether they are on or off because consumers do not act in a way that creates demand on the basis of information regarding the current state of the appliance, nor should they. The demand is a fractional rate between zero and one. The period t for the cycle of an arbitrary thermostatic appliance is given by

$$t = t_{on} + t_{off} = \left(T_{off} - T_{on}\right)\left(\frac{1}{r_{on}} - \frac{1}{r_{off}}\right) \tag{6.4}$$

The signs of r_{on} and r_{off} must be consistent with the respective positions of T_{on} and T_{off}; i.e., heating appliances have $T_{on} < T_{off}$, $r_{on} > 0$, and $r_{off} < 0$ and cooling appliances have $T_{on} > T_{off}$, $r_{on} < 0$, and $r_{off} > 0$. The fundamental period of a thermostatic load depends only on the thermal and control properties of the load. Thus it would be more accurate to assert that the distribution of states depends on the thermal properties, the control parameters, and the demand.

Some authors prefer to use diversity factors, k_d, when describing the aggregate effect of randomly distributed individual loads on distribution circuit loading calculations (Glosh et al. 1997). *IEEE Standard 141-1993* defines diversity factor in §2.4.1.3.5 as follows:

The ratio of the sum of the individual non-coincident maximum demands of various subdivisions of the system to the maximum demand of the complete systems. The (unofficial) term diversity, as distinguished from diversity factor refers to the percent of time available that a machine [...] has its maximum or nominal load or demand [...].

Diversity is defined here as the duty cycle discussed above. For thermostatic devices, the diversity factor is computed as $k_d = t / t_{on}$, which is the inverse of the duty cycle (Lu and Chassin 2004). We refer

---

(a) In the case of cooling appliances, the heating rate is r_{off} and the cooling rate r_{on}.

to duty cycle when discussing the properties of a single load or a class of similar loads, and to diversity and diversity factor when discussing the aggregate behavior of populations of loads.

5.2 Discrepancies of Classical Model

Thermostatic loads are readily modeled using a state space in which two dimensions are considered, temperature and mode (Guttromson et al. 2003; Lu and Chassin 2004; Hatziargyriou et al. 1990). The first version of PDSS implemented an agent-based model of thousands of loads operating independently in response to real-time price signals posted at arbitrary times, which are in turn updated in response to the load. To eliminate the simulation's initial settling time, the initial distribution of the thermostatic devices over the state space must be the steady-state condition. Thus we computed the probability density function for the steady-state condition for an arbitrary thermostatic device. The lack of discussion of steady-state distributions in the load modeling literature suggested to the designers of PDSS that the distribution was generally regarded as obvious and intuitive: if the demand and period were constant, the duty cycle was the only governing factor, and both the steady-state and initial probability density functions for thermostatic devices must be uniform such that

$$\Pr\{x \leq T < x + dx\} =$$

$$
\begin{cases}
x < T_{on} & : & 0 \\
T_{on} \leq x \leq T_{off} & : & \dfrac{dx}{\left|T_{off} - T_{on}\right|} \\
x > T_{off} & : & 0
\end{cases}
\tag{5}
$$

$$\Pr\{on\} = \varphi \quad , \quad \Pr\{off\} = 1 - \varphi$$

However, while testing the first version of PDSS, two problems were observed when the demand was initially non-zero and called into question whether the initial conditions were the same as steady state: 1) an initial transient load swing similar to a response to a price change always occurred, and 2) the steady-state diversity was not the value expected. Several possible causes for this inconsistency were investigated. The most likely included

1. a coding error that introduced incorrect load responses to non-zero demand

2. a thermal or control model that caused improper cycling of the appliances

3. a steady-state distribution that is not uniform.

Causes 1 and 2 were of particular concern because they might contribute to erroneous results from many PDSS analyses, whereas Cause 3 was not particularly serious and would not likely contribute to serious errors. However it is interesting in that it could preclude Causes 1 and 2 if its correction eliminates the initial transient and erroneous diversity prediction and could provide new insight into the behavior of thermostatic loads in general.

As is common in these situations, extensive analysis and testing of the simulation never did completely preclude Causes 1 and 2, but they appeared to be very unlikely. Therefore, Cause 3 was investigated in spite of the fact that it appeared improbable. A simplified version of the PDSS model was
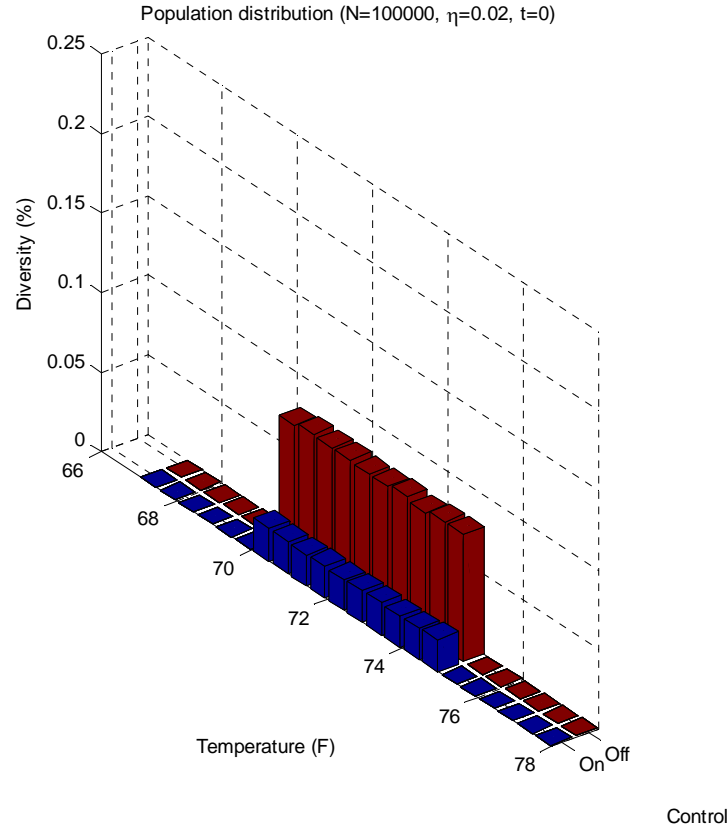
Figure 5.1. The Initial State Distribution is Uniform and Remains Uniform When Demand is Zero

constructed in MATLAB, tested and examined to determine the true steady-state distribution. The initial distribution used by PDSS was provided to the verification simulation and is shown in Figure 5.1. This is a continuous model, much like the one the PDSS agents use, and not like the discrete model used in the consideration of the diversity impact of price-responsive thermostatic end uses (Lu and Chassin 2004).

The result of the simulations shows that the final steady-state distributions differ depending on whether the demand is zero. When the demand is zero, the steady-state distribution remains uniform, just as it was at initial conditions. However, Figure 5.2 reveals a phenomenon that has not been previously described: the steady-state distribution is not uniform when demand is non-zero; rather, it follows what appears to be an exponential curve. This suggests that the original assumption that demand did not affect the steady-state distribution was incorrect.

The second discrepancy that was noted is shown in Figure 5.3. According to the standard model, the diversity is expected to be 0.2. However, when the demand is non-zero, the diversity exceeds the predicted value, suggesting that the load diversity factor is closer to unity than expected based solely on knowledge of the thermal properties of the end uses.
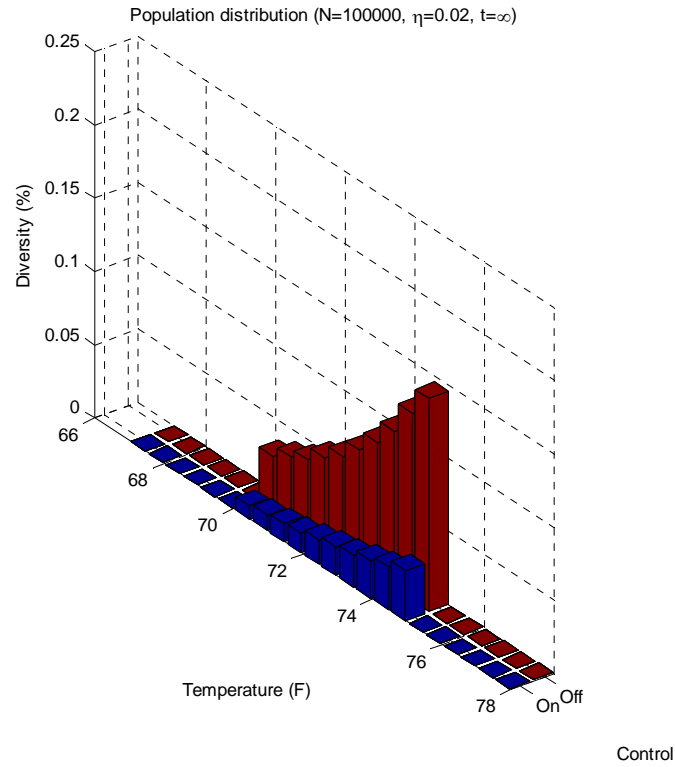
Figure 5.2.   The Steady State Distribution is Not Uniform When Demand is Non-Zero
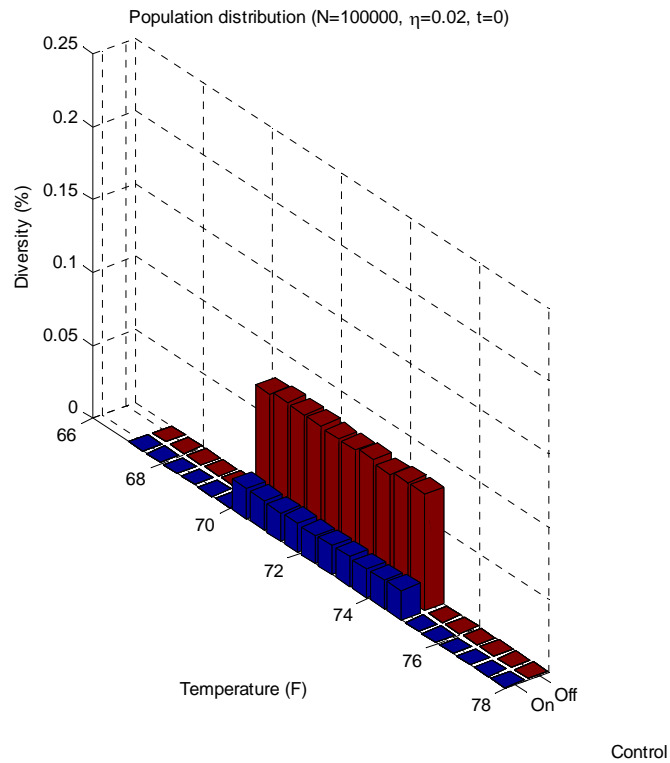
Figure 5.3.   Initial State Distribution is Uniform and Remains Uniform when Demand is Zero

Population distribution (N=100000, η=0.02, t=∞)
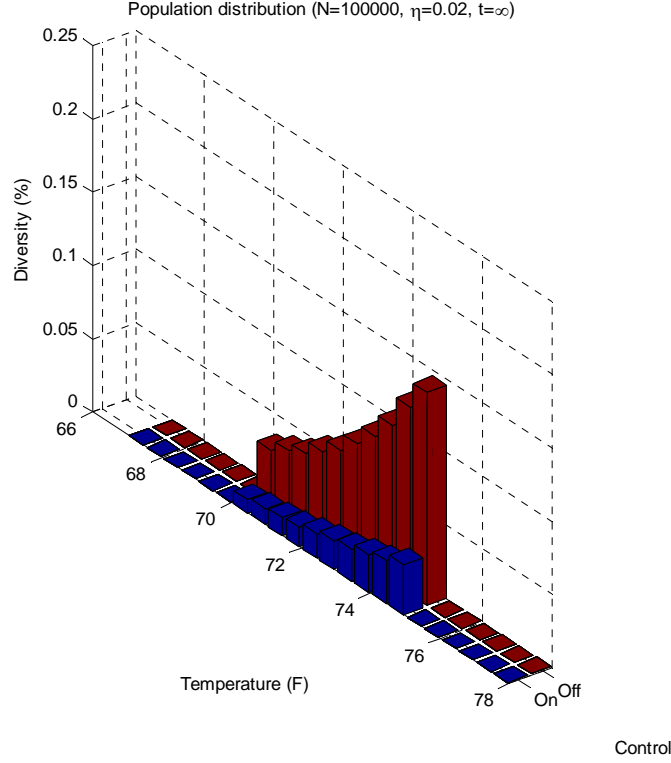
Figure 5.4. Steady-State Distribution Is Not Uniform when Demand Is Non-Zero

5.3 Analysis of State Dynamics

When a thermostatic end-use device is subject to a demand, it seems we must consider the contribution that demand makes to the steady-state dynamics of the end uses. It seems reasonable to consider that as the population of devices in the *off* mode cools down and approaches the *on* set-point, T_{on}, they are uniformly drawn from by demand. Those drawn upon by demand transition immediately to the *on* mode at a temperature somewhere between the current temperature and T_{on}, depending on the end use. The model assumes the temperature remains unchanged. The rate at which devices are drawn from the *off* mode is the demand rate, η. In Figure 5.4, the typical cycle of a thermostatic device with non-zero demand is illustrated. The number of devices *on* at a given temperature T and a given instant t is denoted $n_{on}(T, t)$ with the second argument omitted to emphasize that we are interested in steady-state equilibriums.[a] It is similar for the number $n_{off}(T, t)$ of devices in the *off mode*.

The functions $n_{off}(T,t)$ *and* $n_{on}(T,t)$ are related to one another and to the constants r_{off}, r_{on}, etc., by differential equations, which are themselves determined from difference equations that govern a quantized model. For clarity, let $n_x(T:dT,t)$ denote the number of devices in the mode $x \in \{on, off\}$ with temperatures between T and $T+dT$ at the exact time t. The variation $dn_x(T:dT)$ in the number of x-mode devices between time t and $t+dt$ is caused by the combined effect of demand, of devices leaving the

---

(a) We have two axes of concepts. On one hand, the solutions to the governing differential equations and the solution to the difference equations that give rise to the former equations. On the other hand, we have the time-dependent and steady-state solutions of those equations. It would be better if the notation made the distinction clear.

temperature range T to $T+dT$ due to natural cooling or heating, and of control changes of mode at set-point temperatures. The latter cause is captured in the boundary conditions. The critical quantities to estimate are the fraction of devices changing to the *on*-mode as a result of demand, the fraction of devices leaving the range T to $T+dT$ during a given period of time dt because of cooling or heating alone, and the impact of demand on the latter.

Demand affects the *off* and *on* modes differently; it depopulates the *off*-mode states and repopulates *on*-modes but affects only the instantaneous rate of heating. To be precise, let r_x denote the instantaneous rate at which the temperatures change in mode x when demand is null; its units are in degrees per second (°C/sec) or some equivalent. The demand, η, is the instantaneous probability of a random appliance usage event, independent of the initial mode of the device; its units may be given in sec^{-1}:

$$\eta = \lim_{\Delta_t \to 0} \frac{prob(use \quad between \quad time \quad t \quad to \quad t+\Delta t)}{\Delta_t} \tag{5.6}$$

For now, the given temperature T is at least $2dT$ away from any temperature end point. The distribution of temperatures is assumed to be approximately uniform within a range of dT units, which is important but relatively easy to assert for large populations of machines as $dT \to 0$. During dt time units and in the absence of demand, the devices in the *off* mode see their temperature decrease by $dt \times r_{off}$ degrees. Thus, at the end of period dt, the range of temperatures for the devices accounted for in $n_{off}(T:dT,t)$ at time t will extend to the left, as shown in Figure 5.5, by a proportion of $r_{off} \times dt/dT$. Under the assumption of a local uniform distribution, cooling alone removes approximately $(r_{off}\, dt/dT)$ $n_{off}(T:dT,t)$ devices from the count $n_{off}(T:dT,t)$; it is similar for devices in the *on* mode. The number of devices accounted for by $n_{off}(T:dT,t)$ that switch *on* due to demand in a duration of dt is $\eta\, n_{off}(T:dT,t)\, dt$. All of this is summarized graphically in the figure and in the finite difference equation (5.5).

$$
\begin{aligned}
dn_{off}(T:dT) \;=\; & -\eta.dt.n_{off}(T:dT,t) \\
& -\frac{r_{off}.dt}{dT}(1-\eta.dt).n_{off}(T:dT,t) \\
& +\frac{dt.r_{off}}{dT}(1-\eta.dt).n_{off}(T+dT:dT,t)
\end{aligned}
\tag{5.7}
$$

or, equivalently,

$$
\begin{aligned}
\frac{dn_{off}(T:dT)}{dt} \;=\; & -\eta.n_{off}(T:dT,t) \\
& +r_{off}\frac{n_{off}(T+dT:dT,t)-n_{off}(T:dT,t)}{dT}(1-\eta.dt)
\end{aligned}
\tag{5.8}
$$

A partial differential equation is obtained taking the limit of this equation as both dt and dT tend to zero, namely

$$\frac{\partial n_{off}}{\partial t} = -\eta.n_{off} + r_{off}\frac{\partial n_{off}}{\partial T} \tag{5.9}$$

5.7

Figure 5.5. **Load Diversity at Equilibrium Is Not as Predicted by the Classical Model (dash-dot) when the Demand is Non-Zero. Tick**

The equations for the *on* mode devices, away from the temperature end points, are derived in a similar fashion; the main difference is in the effect of demand on the heating rate. Consider devices in the *on* mode and whose temperature is within T and T+ΔT. The average temperature of those devices is T+½ΔT. At each instant during time *dt*, there are about $\eta\, n_{off}(T{:}\Delta T,t)$ of those devices with a fixed temperature because of demand. The average temperature of those devices after Δt time units is approximated by the sum of integrals:

$$\eta\int_{s=T}^{T+\Delta t} s\, dT + (1-\eta)\int_{T}^{T+\Delta t}\left(s+r_{on}\Delta t\right)dT = T + \frac{\Delta T}{2} + (1-\eta)r_{on}. \tag{5.10}$$

Effectively the instantaneous heating rate under demand is $r_{on}(1-\eta)$. The governing finite difference equations for on-mode devices away from temperature end points are therefore

$$\begin{aligned}
dn_{on}(T:dT) \;=\;& \eta.dt.n_{off}(T:dT,t) \\
& -\frac{r_{off}(1-\eta).dt}{dT}.n_{on}(T:dT,t) \\
& +\frac{r_{off}(1-\eta)dt}{dT}.n_{on}(T-dT:dT,t)
\end{aligned} \tag{5.11}$$

The corresponding governing differential equation for *on*-mode devices is obtained by taking the limit of this equation as *dT* and *dt* approach zero, namely

$$\frac{\partial n_{on}}{\partial t} = \eta . n_{off} - r_{on}(1-\eta)\frac{\partial n_{on}}{\partial T} .$$

(5.12)

5.4   Time Independent Distributions

The time-independent solutions to this problem are obtained by setting to zero the time derivative given in Eq. (5.7) and (5.10), which leads us to a pair of ordinary differential equations in T:

$$\frac{\partial n_{off}}{\partial T} = \frac{\eta}{r_{off}} n_{off}$$

(5.13)

$$\frac{\partial n_{on}}{\partial T} = \frac{\eta}{r_{on}(1-\eta)} n_{off}$$

(5.14)

The time-independent functions n_{off} and n_{on} can be obtained, up to two constants K_{off} and K_{on}, by integrating Eq. (5.11) and (5.12). Note that when $\eta=1$, Eq. (5.14) forces $n_{off} = 0$.

$$n_{off} = K_{off} e^{\frac{\eta}{r_{off}}T}$$

(5.15)

$$n_{off} = K_{off} e^{\frac{\eta}{r_{off}}T}$$

(5.16)

The boundary conditions for the time independent solutions in (5.14) and (5.15) simply states that inputs and outputs must match at the temperature end points, which takes the algebraic form

$$r_{on}(1-\eta)n_{on} = r_{off} n_{off}$$

(5.17)

It follows that $K_{on} = 0$. We also can compute N_{off} and N_{on}, the total number of machines *off* and *on*, respectively, such that $N_{off} + N_{on} = N$, by integrating (5.15) and (5.16) between T_{on} and T_{off}. The constant K_{off} may be obtained from the equation for mass conservation:

$$\int_{T_{min}}^{T_{max}} n_{on} + n_{off} dT = \left(1 + \frac{r_{off}}{r_{on}(1-\eta)}\right)\int_{T_{min}}^{T_{max}} n_{off} dT = N$$

(5.18)

Explicitly, one finds

$$K_{off} = \frac{N\eta(1-\eta)r_{on}}{\left[(1-\eta)r_{on} + r_{off}\right]\left[e^{\frac{\eta}{r_{off}}T_{max}} - e^{\frac{\eta}{r_{off}}T_{min}}\right]} .$$

(5.19)

Finally, we can compute the load diversity by integrating (5.15) to derive a linear relationship between N_{off} and N_{on} and substituting into the definition of φ.
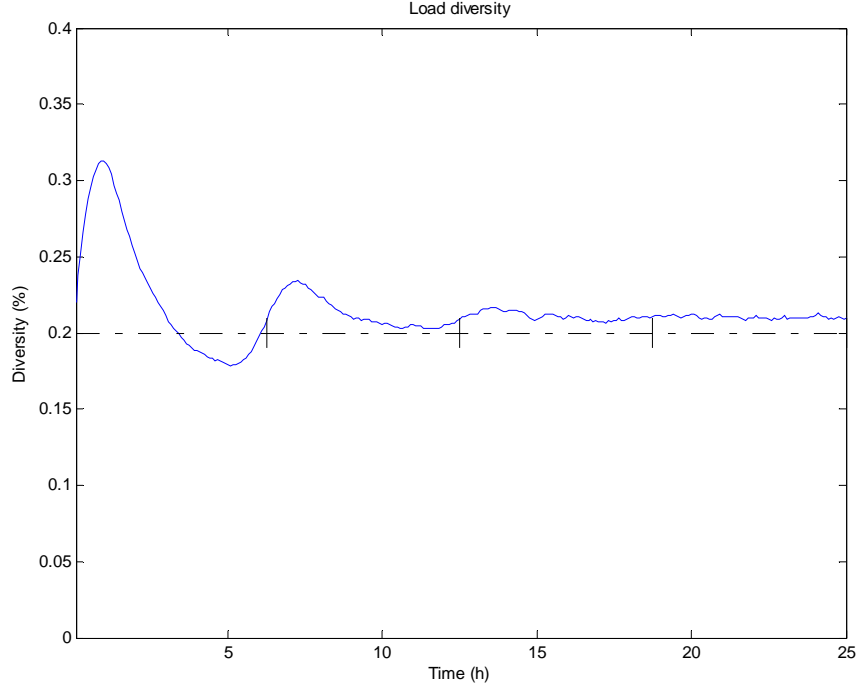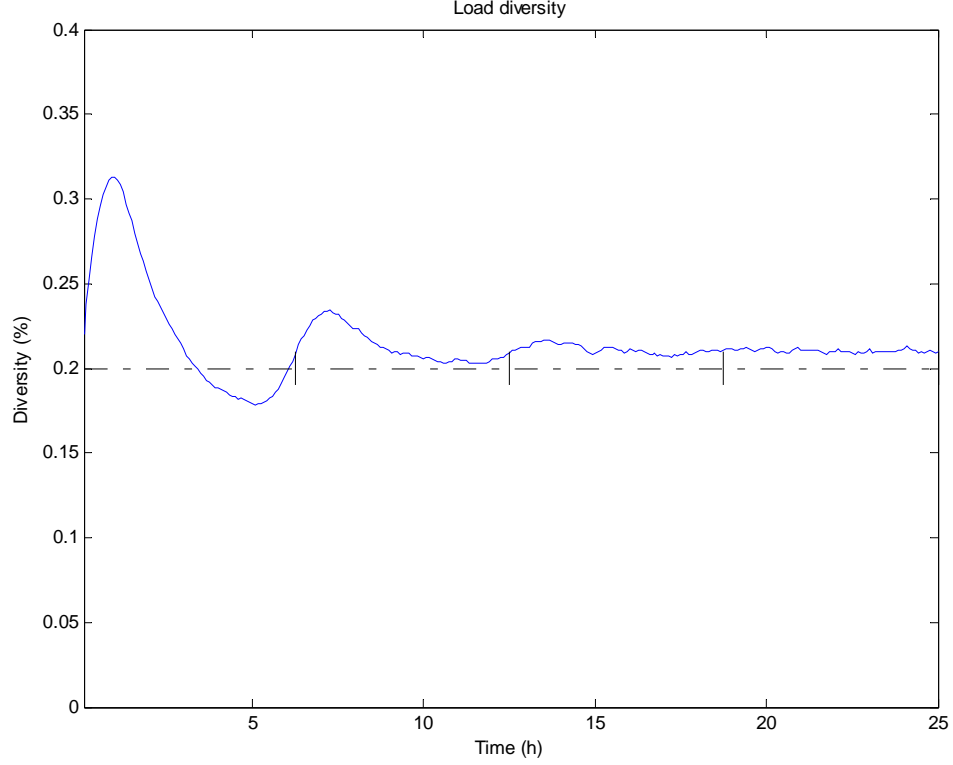
Figure 5.6. **The Load Diversity at Equilibrium is not as Predicted by the Classical Model (dash-dot) when the Demand is Non-Zero. The tick marks indicated the predicted period. Note the transient response to non-equilibrium distribution of states in the initial conditions.**

$$\varphi = \frac{N_{on}}{N_{off} + N_{on}}$$

$$= \frac{{}^{r_{off}}\!\!\diagup\!\!{}_{r_{on}(1-\eta)}}{1 + {}^{r_{off}}\!\!\diagup\!\!{}_{r_{on}(1-\eta)}} \tag{6.20}$$

$$= \frac{r_{off}}{r_{off} + (1-\eta)r_{on}}$$

This result is compared with the diversity observed in the simulation, as shown in Figure 5.6. The deviation of the model from the simulation is one of the principal flaws of numerical simulations of this type. The discrete nature of the simulation can result in variations in the results for certain parameters.

Superficially, this result does not seem to affect load forecasting and distribution system planning to any great extent. As previously noted, the impact of incorrectly computing the steady-state distribution resulted only in an initialization transient in the simulation (a rather common occurrence) and a small error in the expected diversity. However, it has already been shown that price fluctuations can dramatically impact load diversity and that demand plays a role in dampening that phenomenon (Lu and Chassin

5.10
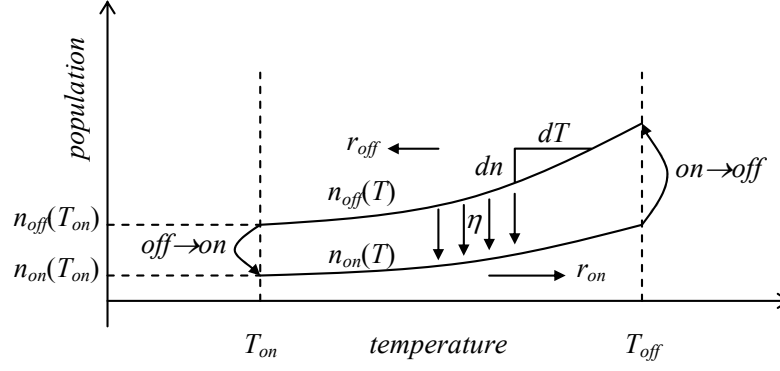
Figure 5.7. **The Thermostatic End-Use Cycle for a Heating Regime With Demand** $\eta > 0$
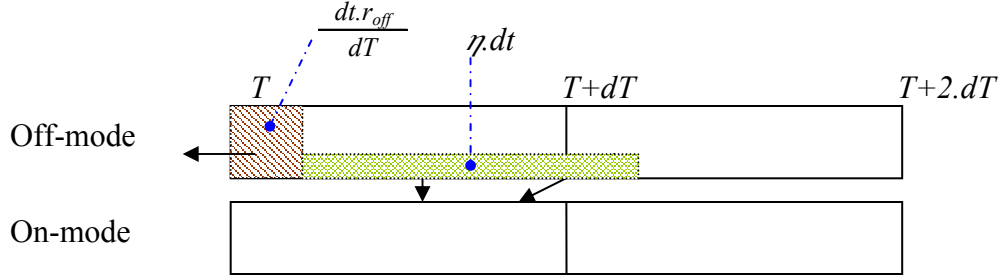
Figure 5.8. **Shaded Areas Represent Devices that Change Their Quantized State Between Time t and t+dt**

2004). Further, we treated demand as a time-independent constant. The natural cycle of the devices under consideration is often similar to the cycle of diurnal demand fluctuations. Thus the above consideration would benefit greatly from including a more general time-dependent demand function, similar to those observed in end-use metered data (Pratt et al. 1989). It is reasonable to believe that such a refinement would lead to much better insights into the relationship between price fluctuation and load diversity.

5.5   Time-Dependent Distributions

First, the only distributions that satisfy the flow equations (5.15) at all temperatures are time-independent. To see this, we let q and Q be the two variables defined implicitly by the relations

$$
\begin{aligned}
t &= -q + Q \\
T &= r_{off}.q + r_{on}(1-\eta)Q
\end{aligned}
$$

(5.21)

We can write (5.9) and (5.12) in terms of q and Q as:
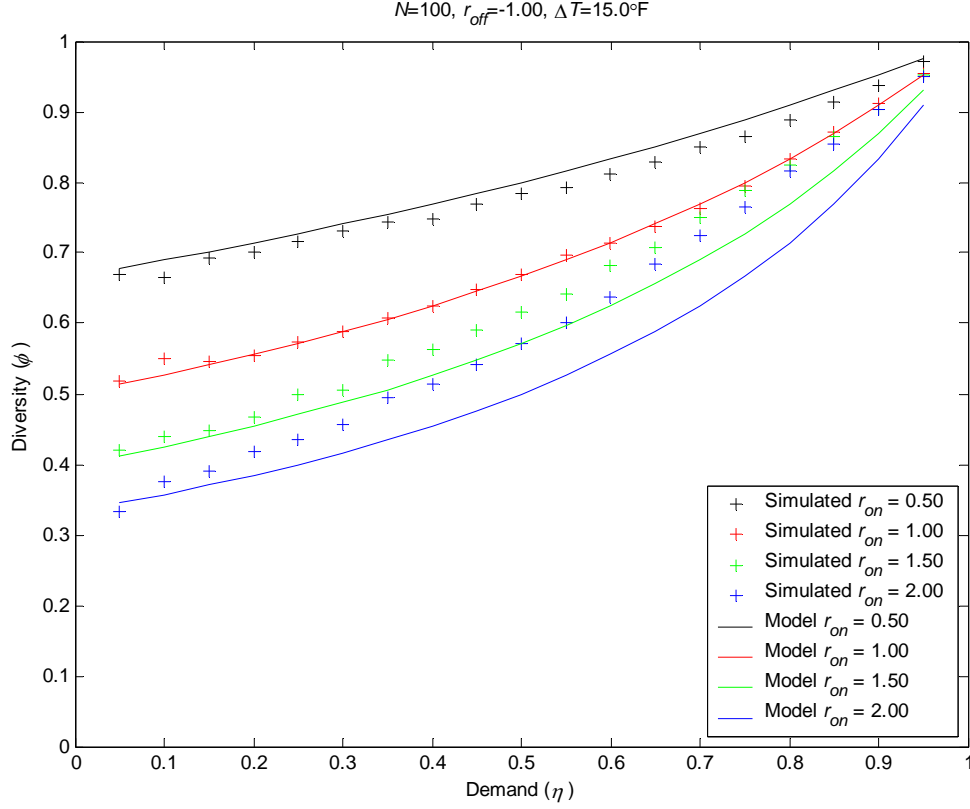
Figure 5.9. **Expected Load Diversity Compared with 100 Simulated Water Heaters with Varying Heating Rates**

$$\frac{\partial n_{off}}{\partial q} = -\frac{\partial n_{off}}{\partial t} + r_{off}\frac{\partial n_{off}}{\partial T} = \eta.n_{off} \tag{5.22}$$

$$\frac{\partial n_{on}}{\partial Q} = \frac{\partial n_{on}}{\partial t} + r_{on}(1-\eta)\frac{\partial n_{on}}{\partial T} = \eta.n_{off} \tag{5.23}$$

Equation (5.22) gives immediately an expression for n_{off}, namely

$$n_{off} = k_{off}(Q).e^{\eta.q} \tag{5.24}$$

where $k_{off}(Q)$ is an unknown function of Q alone. Combining the boundary conditions Eq. (5.15), with Eq. (5.21) and (5.23) gives

$$\frac{\partial n_{on}}{\partial Q} = \frac{r_{off}}{r_{on}(1-\eta)}.e^{\eta q}\frac{\partial k_{off}(Q)}{\partial Q} = \eta.k_{off}(Q)e^{\eta q} \tag{5.25}$$

Considering the second equality, leads to the differential equation

5.12

$$\frac{\partial k_{off}(Q)}{\partial Q} = \eta \frac{r_{on}(1-\eta)}{r_{off}} k_{off}(Q),$$
(5.26)

whose solution up to a constant h is

$$k_{off}(Q) = h.e^{\eta(1-\eta)\frac{r_{on}}{r_{off}}Q}.$$
(5.27)

and can be substituted back into (6.23) to give

$$n_{off}(q,Q) = h.e^{\frac{\eta}{r_{off}}(Q(1-\eta).r_{on}+q.r_{off})},$$
(5.28)

and after substitution of t and T for q and Q

$$n_{off}(t,T) = h.e^{\frac{\eta}{r_{off}}T}.$$
(5.29)

The latter is independent of time and so must be n_{on}.

The standard model of load diversity fails to predict accurately the steady-state diversity observed in highly accurate simulations of end-use load behavior. The discrepancies, when examined closely, suggest that the steady-state distribution of end-use device states is not uniform, as generally assumed. This fact leads to a more accurate expression of the steady-state distribution of end-use device states and can be used to correctly initialize agent-based end-use simulations and accurately compute the expected load diversity for any given level of end-use demand. It is also expected to provide a better understanding of the steady-state regime of end-use loads and ultimately lead to an improved generalized model of the time-dependent behavior of loads responding to changes in demand and set points such as might be implemented in real-time price control strategies.

6.0   Applications in Power Systems Control Strategies

Existing modeling methods and a discrete state queueing (DSQ) model for thermostatic loads are demonstrated in this section. The background of market-based power system operation and the motivation of developing the DSQ model are introduced.

6.1   Applications to Distributed Systems

Deregulation of the electric industry is creating new opportunities and challenges. The challenges include 1) the imbalance between supply and demand because of reduced generation capacity leading to reduced reserves, 2) the potential of market manipulation by exercise of market power as seen in the California market, and 3) the inability to provide reasonably accurate forecasts of short- and long-term loads and electricity prices because of the lack of proper understanding of the dynamic electric markets and the complexity of the system. These challenges can lead to extreme price volatilities, brownouts, and blackouts.

Many solutions to the challenges faced by electric utilities are market-based approaches including better use of distributed energy resources (DERs). DERs include distributed generation, storage, and demand response. The integration of DERs into the power system control framework is part of the evolutionary advances that allow these resources to actively participate in the energy balance equation. The price can be a powerful signal for independent decision making in distributed control strategies. Currently, in response to imbalances in supply and demand in electricity, many public utilities, independent system operators (ISOs), and municipal utility districts have initiated demand response programs. Some of these programs are voluntary while others provide specific financial rewards for curtailing loads and providing demand relief to the grid. Another option to mitigate price volatility is to make loads more responsive to the electricity price, i.e., to make loads elastic. However, loads will not respond to price signals unless there are significant incentives.

Cost-effective solutions to enable "demand responsive" buildings require significant developments in both hardware and software components and infrastructure. Enabling technologies allow customers to interact directly with supply-side agencies to achieve both operational and financial benefits. Another problem with the current approach is high transaction cost. The transaction costs include cost of enabling the loads to be price responsive, the cost of monitoring the loads, and the cost of reconciling and verifying the load reduction for payments. To study the effects of price responsive DERs on the electric power system requires generation and load models that can capture the dynamic coupling between the energy market and the physical operation of the power system in appropriate time frames.

Unless the interaction between the loads and the grid are fully automated, transaction costs cannot be lowered significantly. To automate the interactions we need to better understand how loads react. In addition, to reduce price volatility and improve reliability of the transmission and distribution (T&D) network, there is a need for dynamic pricing of electricity, accurate forecast of short- and long-term aggregate loads and price, improved hardware and software components, and infrastructure. This section addresses some of the issues relating the accurate forecast of loads using the first-principles approach. In this section, we apply some of the principles developed in the earlier sections to better understand the behavior of the loads under different control strategies.

This study on modeling power systems as complex adaptive systems includes applications of control strategies based on price signals that make DERs more demand responsive. To implement various price-responsive control strategies, the behavior of individual loads as well as of aggregated loads is required. Though we described how the statistical mechanics approach can be used to predict aggregate loads and wholesale prices as time-dependent properties of maker-based power systems, the theory could not be fully developed and validated due to resource constraints. Therefore, we had to use alternative modeling approaches to study the benefits of technologies, controls, contracts, and market.

The following discussion explains what makes the electricity market a complex system, the features of the system, and how the fundamentals of system behavior can be modeled to simulate economic operation of the electricity market and physical control over the power systems. Also discussed is using the discrete state queueing model and how it works, and the advantages and the limitations of this model compared with other approaches.

6.2   Background on Modeling the Electricity Market

Figure 6.1 depicts the electricity market and illustrates the interaction of each component. Multiple parties in the bulk power system engage in competitive markets to fulfill their own economic objectives.

In the full transacting bid-based market model, the bulk market operator collects bids from generation companies (GenCos) to obtain a supply curve S and load servicing entities (LSEs) to obtain a demand curve D. The market clears at a price P where S and D intersect. In the real-time market, bidding is done hourly (or sometime less than hourly) and clearing price established hourly.

A healthy market requires perfect competition, in which no producer or consumer has the power to influence prices. The market price is then set around the competitive price, which leads to an efficient market, according to the economic definition of Pareto efficiency (Pindyck and Rubinfield 2000).

In an electricity market, the price dynamic is determined mainly by the system loads and to a smaller extent the generation and transmission availabilities. As the demand exhibits cyclical daily, weekly, and seasonal patterns, so also does the market clearing price. Price volatility (Alvarado 2000) occurs when there is a significant imbalance between the supply and demand caused either by the physical constraints of the generation and transmission capacities or by market power, which usually is exercised by the supply side because of the lack of elasticity on the demand side. Therefore, increased demand elasticity has a moderating effect on suppliers, reducing their ability to exert market power and set the price above the competitive price.

An LSE's bulk market bid includes demand Q at price B, which reflect their arbitrage expectation of the current forward market clearing price P and their forecast of current load Q. The LSE develops the bid to maximize its profit over time and meet demand fully. Implementing price-responsive demand programs extends the control to end-user appliances and makes it possible for LSEs to have intelligent control over the load, increasing the demand elasticity and contributing to moderate price volatility.

The integration of DERs into the power system control framework is one of the evolutionary advances that allow these resources to actively participate in the energy balance equation. An LSE that provides a local market for price-responsive DERs can interact with the bulk market signals by adjusting
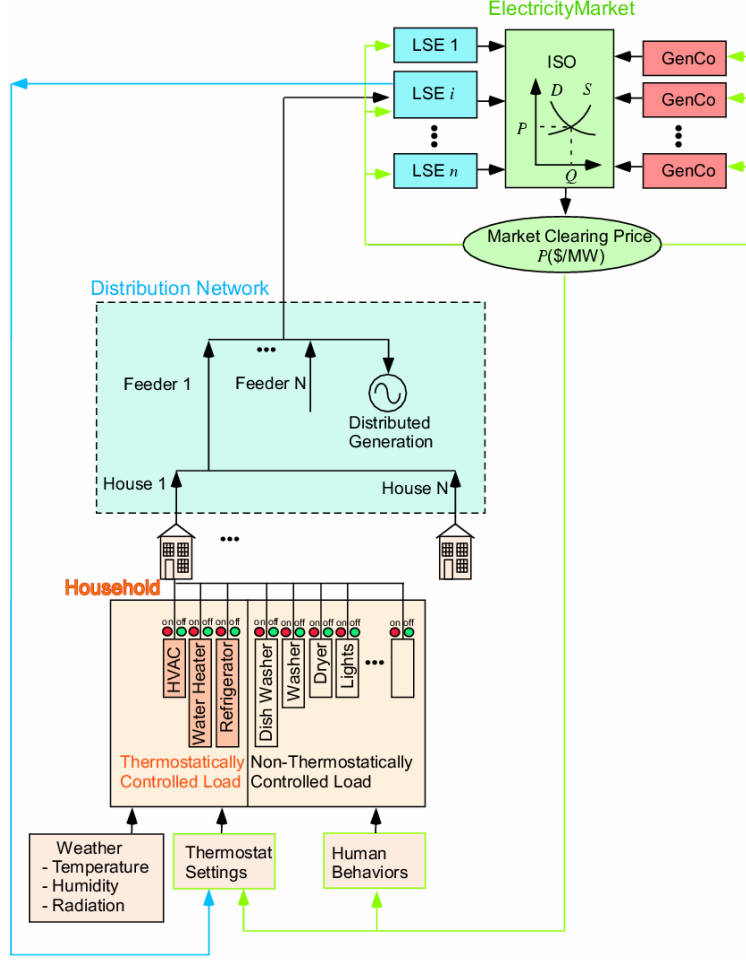
Figure 6.1.   Block Diagram of a Typical Electricity Market

the offer for generation based upon the price sensitivity of his DERs, like the ith LSE shown in Figure 6.2. The bulk market clearing price, P, is a function of all n LSE bids, which can be represented as

$$P = f(Q_1, Q_2, ..., Q_n) \tag{6.1}$$

where Q_i is the supply bid of the ith LSE. Similarly, each LSE bids the expected balance of power at an expected price B_i into the bulk market. The control function of the ith LSE is

$$Q_i = g_i(B_i) \tag{6.2}$$

In principle, the control objective of an LSE is to maximize profit over time period T. However, it is typical for today's LSE to consider load as a given and thus seek only to minimize their costs, C, reducing the optimization problem:

$$\min\left[\int_0^T Q_i(t)P(t)dt\right] \tag{6.3}$$
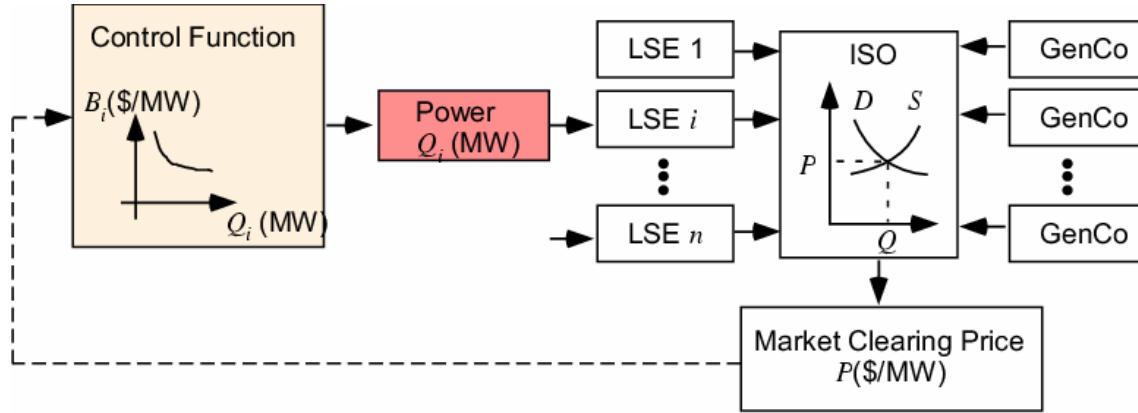
Figure 6.2.   Block Diagram of a Typical Multilayer Electricity Market

If distribution generation is considered in the payment optimization by contracting directly with the LSEs, the optimization problem can be reformulated as

$$\min(\sum_{t=0}^{T}(Q_i^w(t)P_w(t) + Q_i^c(t)P_c(t)) \tag{6.4}$$

where P_w is the wholesale market price, P_c is the contract price with the distribution generators, Q_i^w is the power bid into the wholesale market, and Q_i^c is the power contracted with the distributed generators. So far, the simulations of the market-based LSE demand response programs seem to be a matter of solving an optimization problem with constraints. So where is the complexity? If an LSE can forecast the market clearing price, P, and its load, Q_i, in a deterministic manner, the whole system would not be a complex one. Unfortunately, as shown in Figure 6.2, the exogenous inputs such as weather, human behaviors, and controls (for example, thermostat settings) constantly influence the individual load outputs, the effects of which can be nonlinear. Furthermore, the feedback control loops nested in appliance control mechanisms also affect the behavior of the load. To further complicate the problem, thermostatically controlled appliances (TCAs) have a thermal capacitance that needs to be tracked down to determine how long they can stay off-line.

Bar-Yam (2000) notes that the study of complex systems takes two approaches: modeling the relationship between elements and parts and reconstructing the whole, and moving from an understanding of the relationship of systems to their descriptions. To model the electricity market successfully, it is essential to model the elements and parts and reconstruct the system. This is a bottom-up approach. By averaging the properties of each element (each household), we can derive the properties of the complex system (the electricity market) as a whole. Because the study of complex systems is about understanding the relationship of the parts to the system as a whole, many statistical tools are useful. Queueing theory and statistical mechanics are the two major tools used in this research. The next section reviews the current modeling method and explains why the state queueing model was selected as our approach to simulating market-based demand response programs.

7.0   Modeling

Residential appliances are categorized into those that are thermostatically controlled and those that are not. There are three major TCAs in a residence: HVAC systems, electric water heaters, and refrigerators. The energy consumption of the TCAs can be estimated using simplified first-principles models such as an equivalent thermal parameter (Pratt and Taylor 1994). This approach models the cooling and heating loads in the residence or small commercial building as a function of a few lumped parameters (effective envelope conductance, effective thermal mass, and effective solar apertures) and COPs, weather, internal gains, and thermostat set points. The loads from hot water heaters, refrigerators, and freezers can also be modeled using this simplified approach.

7.1   Electric Hot Water Heater

The thermal model used to simulate TCAs for this study is based on an Equivalent Thermal Parameter (ETP) approach (Pratt and Taylor 1994). Figure 7.1 shows an ETP representation of an electric water heater. The heat transfer properties are represented by equivalent circuits with associated parameters for this type of TCA. Because their characteristics are similar to other TCAs, we focused on water heater models for illustrative purposes. The other types of TCAs are modeled with similar heat transfer models.

Figure 7.2 shows the thermal behavior (temperature of stored water) of a water heater unit over time (assuming there is no water draw). The rising curves indicate the water heater is *on*, and the falling curves represent the standby (or cooling down) periods when the heater is *off*. As the water heater unit cycles, the water temperature in the tank rises and falls accordingly.

The upper and lower limits represent the dead band of the thermostat around the set point. Changing the set point allows regulation of the power consumption of the water heater. To simplify the analysis, a linear approximation of TCA thermal characteristics (Figure 7.3) is used in our model. Because the asymptotic equilibrium temperature (temperature decay when power is shut off to the water heater and the water temperature allowed to coast) is generally much lower (ambient temperature), the exponential falling curve is nearly linear between the upper limit and the lower limit, as shown in Figure 7.3.

Similarly, the exponential rising curve can be assumed to be linear between the upper and lower limits because the heat rate is significantly higher than the loss rate. *T+* and *T–* are the upper and lower temperature limit for a given set point *T*. The unit has a cycling time of τ, with an *on* period of τ_{on} and an *off* period of τ_{off}.

Non-TCAs include dishwashers, washers/dryers, and the like. Non-TCA loads are driven by human behavior. For this we use a technique of load deferral that is based on probability distributions. Given a period of increased prices, this approach computes the amount of load deferred for non-thermal loads. The calculation is analogous to a laundry system. Let the probability of doing a load of laundry follow a PDF curve, as well as the amount of laundry in the laundry basket (which represents the queue). Let the queue increase by a linear function (in the absence of better information), and let it drop each time when a load of laundry is done. A form of the Monte Carlo method is used to determine if and when laundry will be done. A new curve is created called the "likelihood" curve. This function is the combination of the
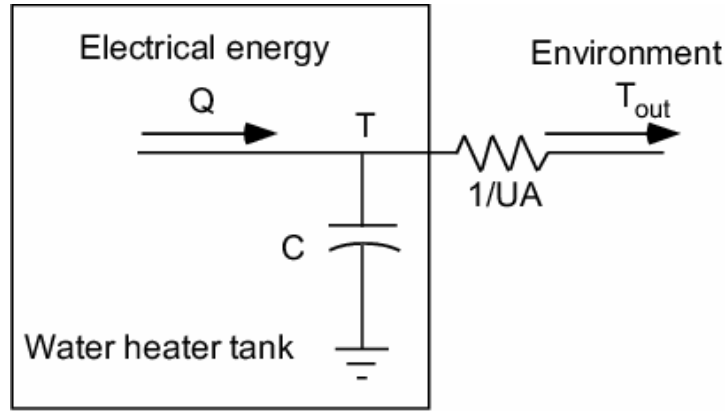
Figure 7.1.   ETP Model for Storage Water Heater
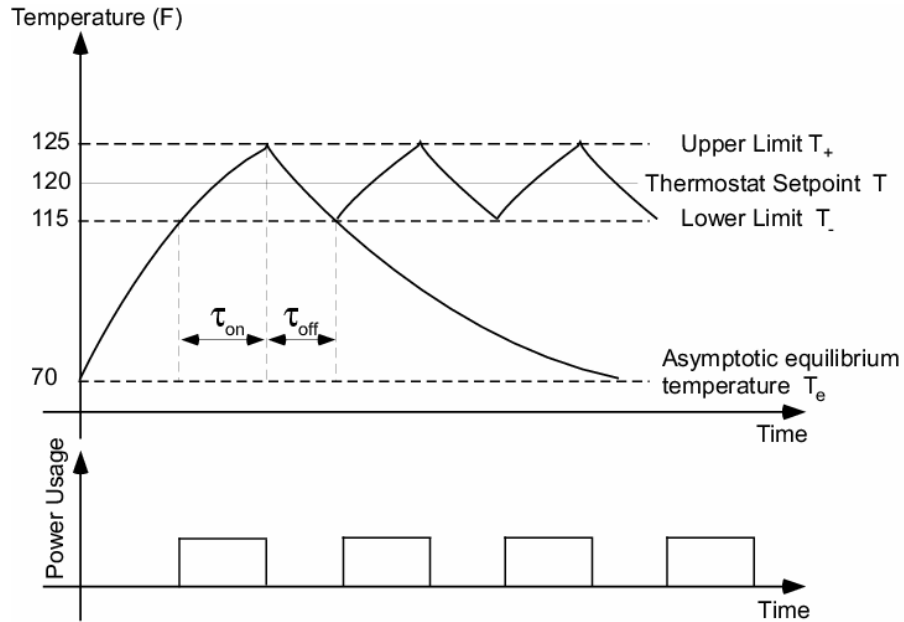
Figure 7.2.   Thermal Behavior of a Typical Water Heater

original PDF curve and the variable queue. The resulting PDF shape and amplitude represent the likelihood of doing a load of laundry on any particular day (see Guttromson et al. (2003) for details).

Non-TCA loads only have two states, *on* and *off*, as shown in Figure 7.4. P_{off} and P_{on} are functions of the control objectives.
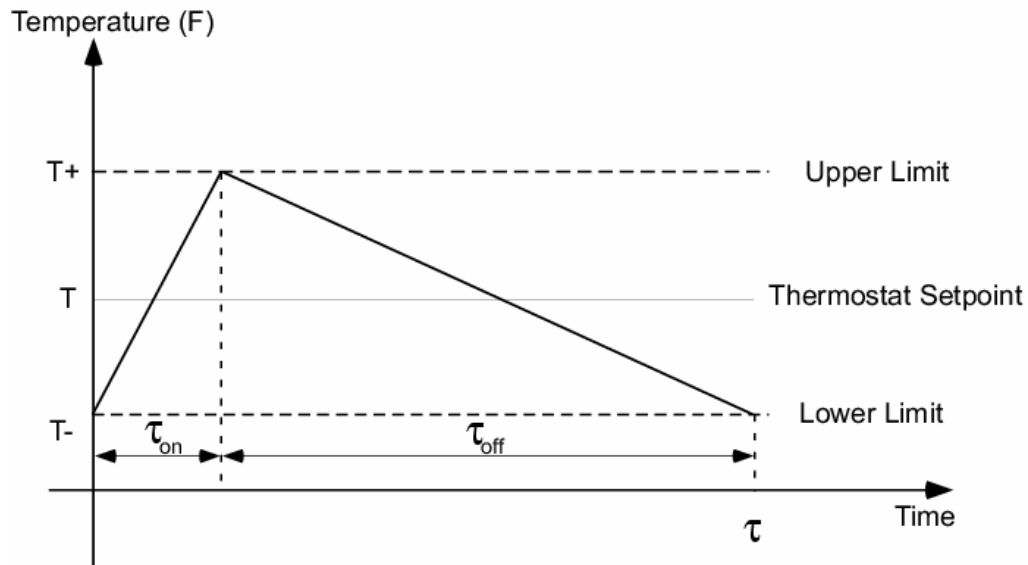
Figure 7.3.   Simplified Thermal Characteristic Curve of a TCA
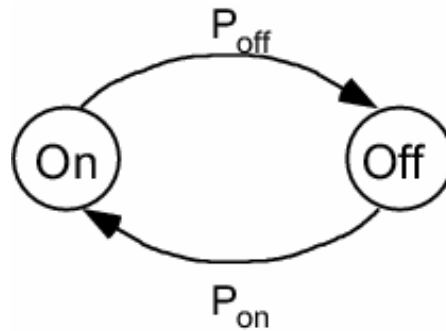
Figure 7.4.   State Queueing Model for Non-TCA Loads

7.2   ETP Model for HVAC System

An ETP model that was used to simulate heating and cooling loads in a typical residence is shown in Figure 7.5. The heat transfer properties are represented by equivalent circuits with associated parameters for modeling the thermostatically controlled HVAC system.

7.3
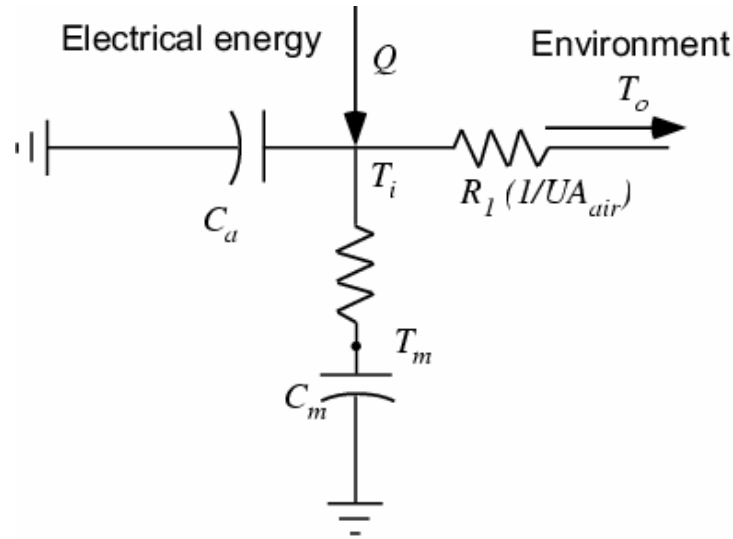
Figure 7.5.   ETP Model for Modeling Residential HVAC Systems

where

C_a = air heat capacity (Btu/°F)

C_m = mass heat capacity (Btu/°F)

Q = heat rate for HVAC (Btu/hr)

UA = the standby heat loss coefficient (Btu/°F·hr)

R_1 = $1/UA_{air}$

R_2 = $1/UA_{mass}$

T_o = Ambient temperature (°F)

T_i = Air temperature inside the house (°F)

T_m = Mass temperature inside the house (°F)

The state-variable description of this ETP model is

$$\dot{x} = Ax + Bu$$
$$y = Cx + Du$$

(7.1)

$$\dot{x} = \begin{bmatrix} \dot{T_i} \\ \dot{T_m} \end{bmatrix} \qquad x = \begin{bmatrix} T_i \\ T_m \end{bmatrix} \qquad u = 1$$

(7.2)

$$A = \begin{bmatrix} -(\dfrac{1}{R_2 C_a} + \dfrac{1}{R_1 C_a}) & \dfrac{1}{R_2 C_a} \\ \dfrac{1}{R_2 C_m} & -\dfrac{1}{R_2 C_m} \end{bmatrix} \quad B = \begin{bmatrix} \dfrac{T_0}{R_1 C_a} + \dfrac{Q}{C_a} \\ 0 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

(7.3)

Solving the model, we can obtain the cooling and heating load and thereby the energy consumed to maintain the residence at the desired comfort level, as shown in Figure 7.6. The rising curves indicate

7.4

Coefficients for T(t) are allowed to change between any solution set (defined by reaching a temperature limit or time step limit)
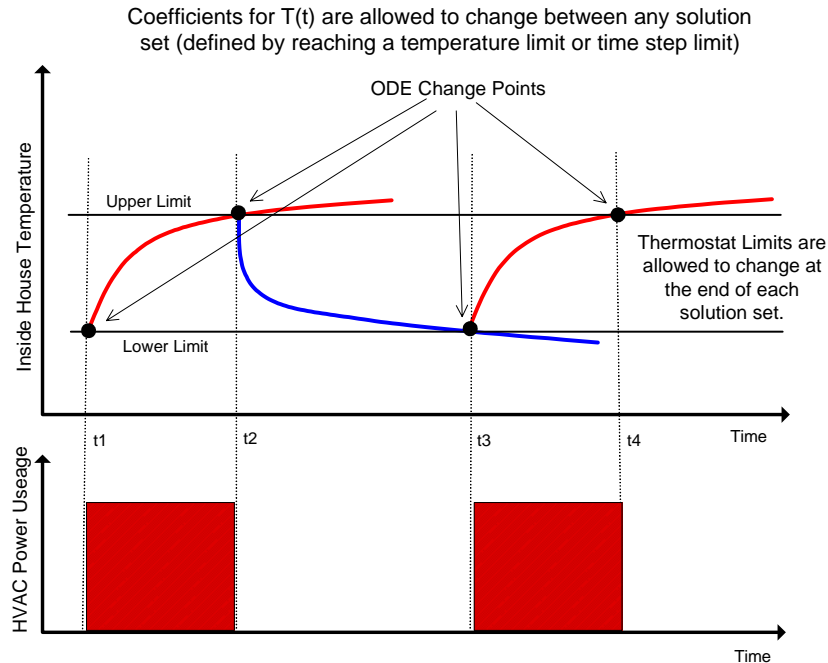
Figure 7.6.   Thermal Behavior of an HVAC System (Guttromson et al. 2003)

operation of the HVAC heater, and show a temperature rise within the house. The falling curve represents a cool down period where the home heating system is not active. As the HVAC unit cycles, the temperature in the house rises and falls with it.

The HVAC load has the following characteristics:

1. The cycling period varies significantly with ambient temperature (Figure 7.6). As the temperature, humidity, internal loads from occupants and other home appliances, and radiation changes, the HVAC load changes.

2. The load profiles vary seasonally. Because the weather changes seasonally, the HVAC load varies on a seasonal pattern. In summer, the house experiences cooling load and in winter it experiences heating load. Typical winter and summer load profiles are shown in Figure 7.7.

3. The load peaks coincide with price peaks in summer but not in winter (Figure 7.8). In western US system where Bonneville collected end-use load data were collected (Pratt et al. 1989), the winter HVAC peaks occurred in early morning while the price peaks usually occur in late afternoon.
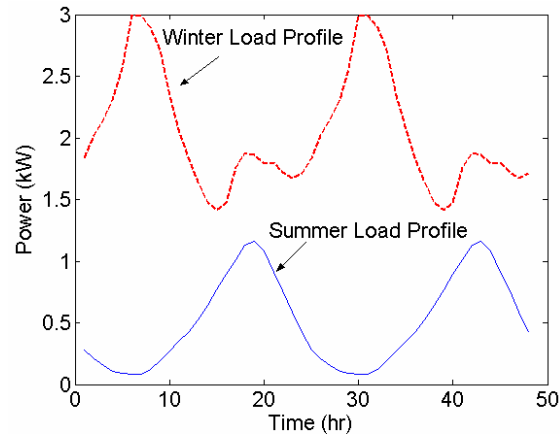
Figure 7.7.   Typical Winter and Summer Load Profiles
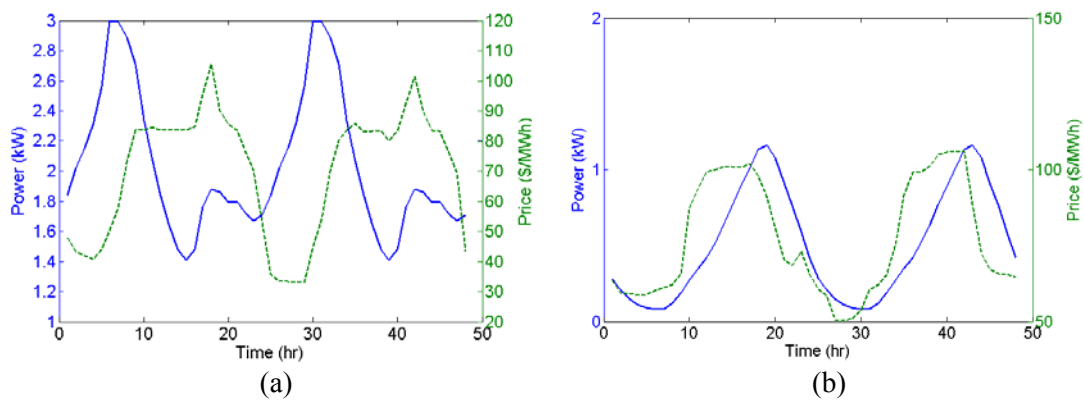
Figure 7.8.   (a) Winter HVAC Load Profile and the DA Market Price Profile, (b) Summer HVAC Load Profile and the DA Market Price Profile

The ASHRAE handbook recommended ranges of thermal comfort in winter and summer are shown in Figure 7.9. In winter, the thermal comfort range is 69°–76°F. In summer, the range is 74°–81°F. In general, people do not want to stay in an environment that is too hot or too cold for several hours. Therefore, they are less likely to implement load curtailment that will cause the HVAC systems to shut down for hours in cold winter or hot summer days.

7.3   Aggregate Load Modeling

There are accurate physical and empirical models that describe the time-variant behavior of an individual load. However, there is not only a need for accurate models of individual loads, there is also a need to accurately model the aggregate load at several different aggregation points in the transmission and distribution network (feeder, distribution substation, transmission substation, control zones, and system level). It is possible to model individual loads and aggregate them to higher levels provided the number of individual loads is manageable (10,000 or less). As the number of loads increases, the complexity of modeling increases. Therefore, there is a need to find alternative modeling techniques that are accurate, efficient, and based on first principles.
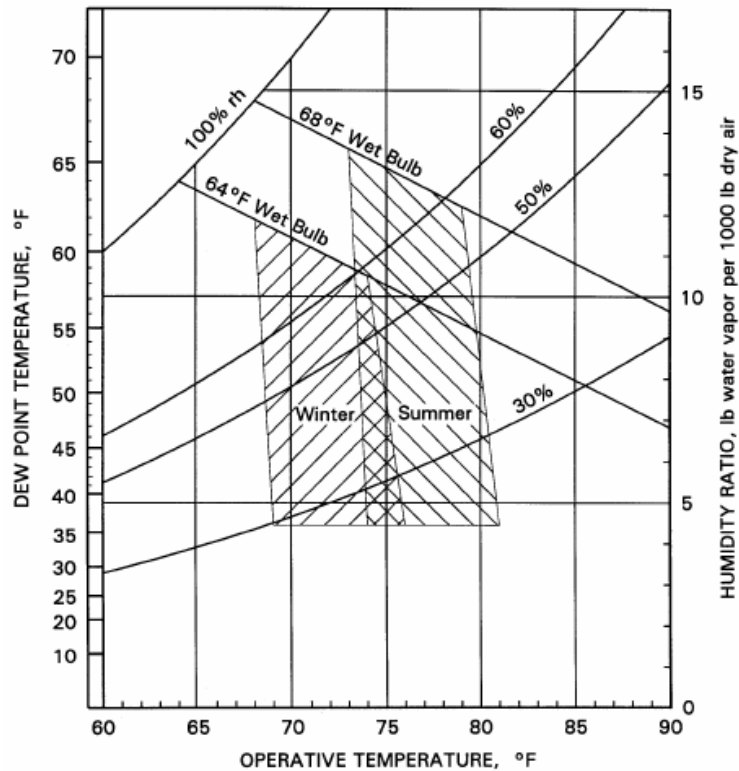
Figure 7.9. **ASHRAE Summer and Winter Comfort Zones from Chapter 8, pp. 8.7; acceptable ranges of temperature and humidity for people in typical summer and winter clothing during primarily sedentary activity. (ASHRAE 2001)**

The objective of this work is to develop an alternative first-principle-based modeling approach that has a mechanism to predict aggregated loads and deal gracefully with the complexity of the problem. Statistical mechanics has been used to model aggregate behavior of complex systems. Although statistical mechanics is a branch of physics developed to answer fundamental questions in physics and chemistry, it has been applied to market-based problems (Schaden 2002; Foley 2002). Statistical mechanics combines quantum mechanics and thermodynamics. Thermodynamics describes properties (energy for example) in macroscopic systems, while quantum mechanics deals with the laws that govern the microscopic behavior of matter. Therefore, statistical mechanics deals with microscopic and macroscopic phenomena and is well suited to aggregate the microscopic behavior of individual loads into a macroscopic level.

7.4   Existing Modeling Methods

Load models developed for TCAs based on statistical analysis (Gellings and Taylor 1981; Pahwa and Brice 1985; Ihara and Schweppe 1981) of historical data cannot model the power output of a TCA when the set point is controlled in response to market prices. Probabilistic calculations of aggregate storage loads (Hatziargyriou 1990) and duty cycle approach (Ryan et al. 1989) have been developed for aggregate cycling load models. These models are based on the end-use data analysis. Historical data collected at the substation level are used to account for non-engineering factors such as weather patterns and customer behaviors. The drawback of these models is that they are empirically driven and noninteractive.
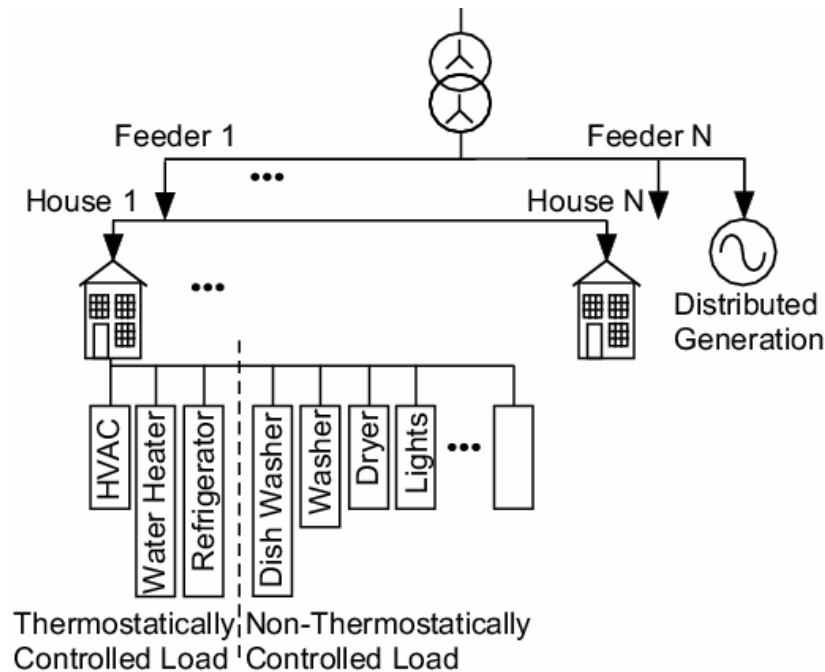
Figure 7.10. A Bottom-Up End-Use Load Aggregation Approach

First-principles or physics-based methodologies for predicting hourly residential HVAC loads can be used to evaluate the impacts of load management programs. Chan et al. (1981) used the diversified load from a limited number of load shapes of individual households to form a load-diversification model. This method takes into account the thermal characteristic of the appliances and leads to a bottom-up approach. First-principles-based load modeling methodologies have been widely used because they can predict the individual load dynamic response and make it possible to obtain the aggregated response of these loads with reasonable accuracy (Molina 2003).

7.4.1   ETP Approach

The end-uses in residential and small commercial buildings can be accurately estimated using simplified first principles models such as the ETP approach (Pratt and Taylor 1994; Taylor and Pratt 1988). An illustrative example is presented in Section 7.4.2. This approach models the cooling and heating loads in the residence or small commercial building as function of few lumped parameters (effective envelop conductance, effective thermal capacitance and effective solar apertures). The loads from hot water heaters, refrigerators and freezers can also be modeled using the ETP approach. ETP modeling approach has been validated and tested extensively. In addition to modeling heating and cooling loads and other thermostatically controlled loads, there is also a need to model discrete events associated with non-TCAs and miscellaneous plug loads that create electric loads on the grid.

The load model that was used to study the price responsive residential load is based on the ETP approach, which takes a bottom-up approach in which detailed physics-based models are used to predict each type of appliance load. Individual loads of each appliance, load on each home, and aggregate load on a group of homes can be estimated using the model shown in Figure 7.10. This approach has many advantages but is computer intensive and does not scale well as the loads grow large (>1,000).
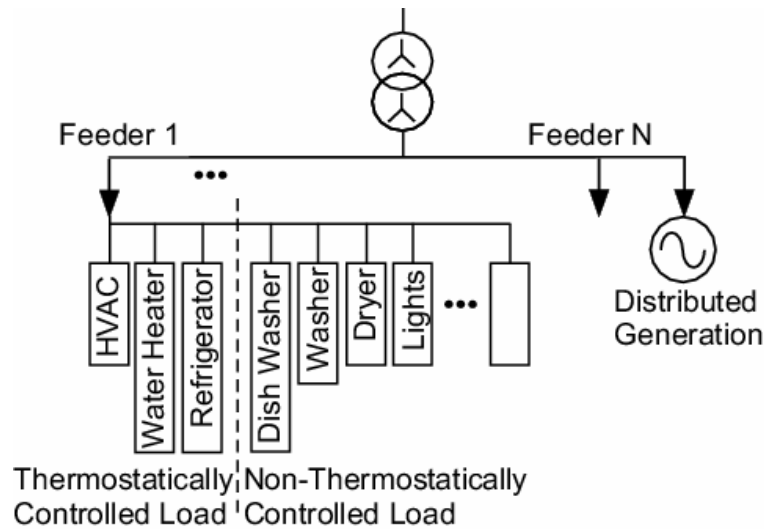
Figure 7.11.   A Feeder-Level Aggregated Load Modeling Approach

7.4.2   SQ Modeling Approach

An aggregated load model using a state queueing model (SQ) approach (Lu et al. 2004a) was developed based on modeling group appliance behaviors to improve the computation speed (Figure 7.11). The SQ model serves as a simplified reduced order feeder equivalent model to the detailed ETP model or the models based on statistical mechanics.

The SQ model differs from the first-principles based approach because it assumes a uniformly randomized load as an initial condition from which the probability distribution function (PDF) is derived. Based on the PDF, the resulting load behaviors are predicted. A computationally efficient state queueing model was developed to model the price response of TCAs. Because TCAs contribute a large fraction of the total residential load, this research can be applied to create simulation tools to study the effectiveness of residential load response programs, which are important in establishing sustainable and successful demand side management (DSM) programs. In the rest of the section the basic concepts of the discrete state queueing model and the simplification and assumption made to set up such models to simulate residential load at distribution feeder level are discussed.

7.4.3   Basic Queueing Models

Queueing theory (Gross and Harris 1998) was developed to provide models to predict the behavior of systems that attempt to provide service for randomly arising demands. Some examples are traffic flow (vehicles, aircraft, people, communications), scheduling (patients in hospitals, jobs on machines, programs on a computer), and facility design (banks, post offices, amusement parks, fast-food restaurants). The discrete state queueing model fundamentally is a cyclic queue. The price response of the thermostatically controlled load results in a discontinuity in the traditional queueing simulation. Our innovation here is that we developed a set of transition rules to rearrange the queue when the price response happens.
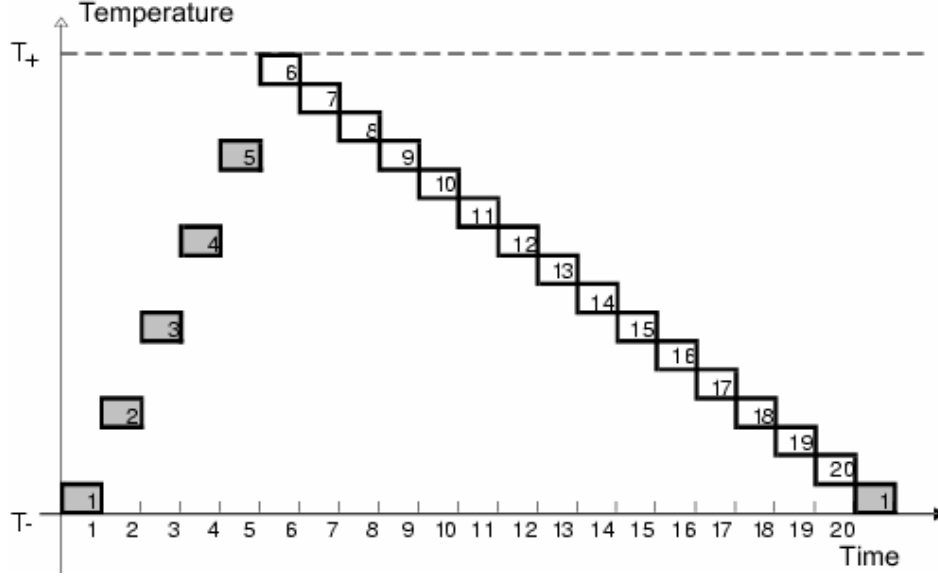
Figure 7.12. SQ Model for Aggregated Water Heater Load

7.4.4 Queueing Model for Hot Water Heaters

A SQ model describes a water heater behavior by tracking the state of the water heater over time. Consider a system containing N water heater units with the initial thermal states shown in Figure 7.12. As an illustration, consider a group of water heater units, whose *on* times are 5 minutes and *off* times are 15 minutes at a given ambient temperature and given set point. Note that in practice, water heaters

usually have cycling times up to 10 hours (when there is no water draw). T_+ and T_- are the upper and lower temperature limits for a given set point T. A state is then defined by both the water temperature in the tank and the on/off status of a unit. We divide the time cycle into 20 states of equal duration such that there are 5 distinct *on* states (shown by shaded boxes in Figure 7.12) and 15 distinct *off* states in a temperature range of $[T_-, T_+]$.

Initially, we assume a uniformly diversified load, the units distributed uniformly among all 20 states. If the whole time period is divided into 20 time steps, at time step 1 we have a nearly equal number of units in each state. At the end of each time step, units move one state ahead. For example, at time step 2, units in State 1 move to State 2, units in State 2 to State 3, and units in State 20 to State 1 (see Table 7.1).

Assuming the average power consumption of each unit is P_{ave}, the total load at each time step is simply the aggregated power output of all the *on* units in States 1 through 5. When the units are uniformly distributed among the states, the expected load P_L can be calculated as

$$E(P_L) = E(N_{on})P_{ave} = \frac{n_{on}}{n} NP_{ave} = \frac{5}{20} NP_{ave} \qquad (7.4)$$

where n is the total number of states, n_{on} is the number of "on" states, N is the number of units, and N_{on} is the number of *on* units.

7.10

Table 7.1.   State Distribution of Water Heater Units

Time \ States	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1
2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2
3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3
4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4
5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5
6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6
7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7
8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8
9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9
10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11
12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12
13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13
14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14
15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20

time →

Table 7.1 shows the state evolution along the timeline. Rows represent time steps and columns represent the unit states. The shaded states are the "on" states. One can calculate the aggregated output by summing up the power consumption of all the units in the *on* states at each time step. For example, if there is a 1-kW unit in each of the 20 states, the aggregated consumption at any time is 5 kW because at any time, there are 5 *on* states in the system.

7.4.5   State Distribution in Response to Price Increase

Let us assume that the TCAs respond to a price increase by lowering their set points (Figure 7.13), which results in a lower on/off temperature limit. A close examination of the moment when the change happens reveals a degeneracy of states: States 4 and 5 combine with States 12 and 9. Because the units in States 4 and 5 immediately turn off as a result of the shift-down of the "off" temperature threshold, they are now at the same temperature and are indistinguishable from States 12 and 9. Furthermore, under the new set point, if we redefine the states in the same pattern used for the original control band, then an "out-of-regime" state transition occurs when the set point changes. For example, the previous States 1 and 2 now become States 4′ and 5′. However, as shown in Figure 7.13, previous States 6 through 14 have to decay into the new state regime because the temperatures of the units are higher than the temperature range set by the new set point; therefore, placing them outside the new regime of states. The temperatures of the units in these "out-of-regime" states will eventually decay into the new temperature setting range shown in Table 2.2. Because the transition of these out-of-regime states into allowed states is deferred, the overall load profile will experience a transient reduction, the length of which is determined by the temperature decay rate. After all state degeneracy and phase alignments are complete; the system will reach a steady state, which is also shown in Table 2.2. If there is one unit in each state prior to the set point change, then the number of units that are *on* at each time step is shown in Figure 7.13.
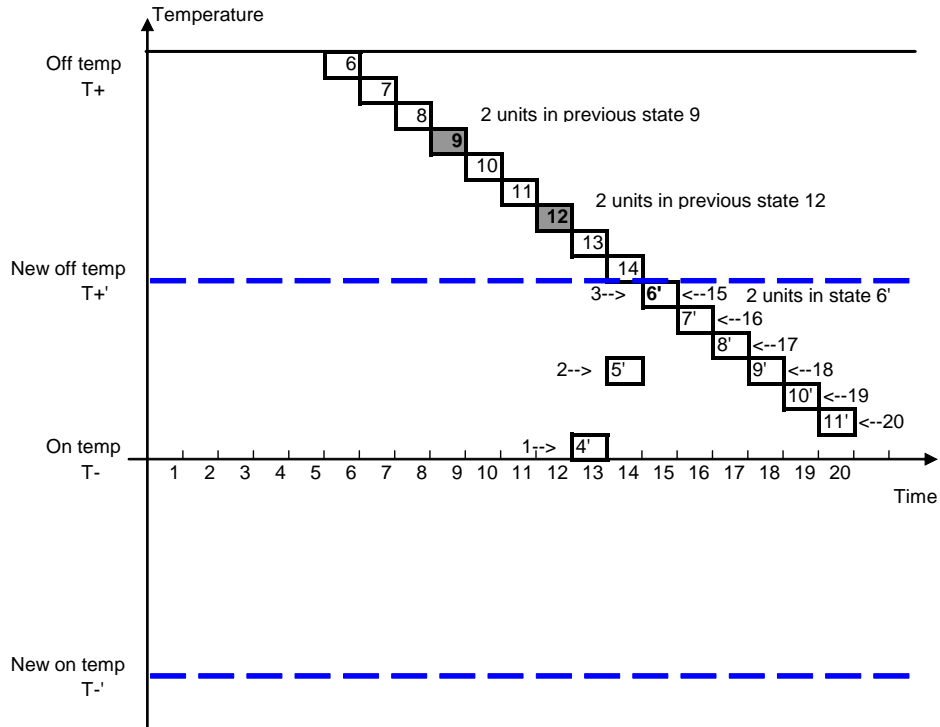
Temperature

Off temp
T+

6
7
8
9 — 2 units in previous state 9
10
11
12 — 2 units in previous state 12
13
14

New off temp
T+'

3--> 6' <--15 — 2 units in state 6'
7' <--16
8' <--17
2--> 5'   9' <--18
10' <--19
11' <--20

On temp
T-

1--> 4'

Time
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

New on temp
T-'

Figure 7.13.   State Degeneracy Representation

Table 7.2.   State Distribution of TCAs after a Price Increase

| State	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	# of "on"
1	4	5	6	off	off	off	off	off	off	off	off	off	off	off	6	7	8	9	10	11	2
2	5	6	7	off	off	off	off	off	off	off	off	off	off	6	7	8	9	10	11	12	1
3	6	7	8	off	off	off	off	off	off	off	off	off	6	7	8	9	10	11	12	13	0
4	7	8	9	6	off	off	off	off	off	off	off	6	7	8	9	10	11	12	13	14	0
5	8	9	10	7	off	off	off	off	off	off	6	7	8	9	10	11	12	13	14	15	0
6	9	10	11	8	off	off	off	off	off	6	7	8	9	10	11	12	13	14	15	16	0
7	10	11	12	9	6	off	off	off	6	7	8	9	10	11	12	13	14	15	16	17	0
8	11	12	13	10	7	off	off	6	7	8	9	10	11	12	13	14	15	16	17	18	0
9	12	13	14	11	8	off	6	7	8	9	10	11	12	13	14	15	16	17	18	19	0
10	13	14	15	12	9	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	0
11	14	15	16	13	10	7	8	9	10	11	12	13	14	15	16	17	18	19	20	1	1
12	15	16	17	14	11	8	9	10	11	12	13	14	15	16	17	18	19	20	1	2	2
13	16	17	18	15	12	9	10	11	12	13	14	15	16	17	18	19	20	1	2	3	3
14	17	18	19	16	13	10	11	12	13	14	15	16	17	18	19	20	1	2	3	4	4
15	18	19	20	17	14	11	12	13	14	15	16	17	18	19	20	1	2	3	4	5	5
16	19	20	1	18	15	12	13	14	15	16	17	18	19	20	1	2	3	4	5	6	6
17	20	1	2	19	16	13	14	15	16	17	18	19	20	1	2	3	4	5	6	7	7
18	1	2	3	20	17	14	15	16	17	18	19	20	1	2	3	4	5	6	7	8	8
19	2	3	4	1	18	15	16	17	18	19	20	1	2	3	4	5	6	7	8	9	9
20	3	4	5	2	19	16	17	18	19	20	1	2	3	4	5	6	7	8	9	10	9
21	4	5	6	3	20	17	18	19	20	1	2	3	4	5	6	7	8	9	10	11	8
22	5	6	7	4	1	18	19	20	1	2	3	4	5	6	7	8	9	10	11	12	8
23	6	7	8	5	2	19	20	1	2	3	4	5	6	7	8	9	10	11	12	13	7
24	7	8	9	6	3	20	1	2	3	4	5	6	7	8	9	10	11	12	13	14	6
25	8	9	10	7	4	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	6
26	9	10	11	8	5	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	5
27	10	11	12	9	6	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	3
28	11	12	13	10	7	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	2
29	12	13	14	11	8	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	1
30	13	14	15	12	9	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	0
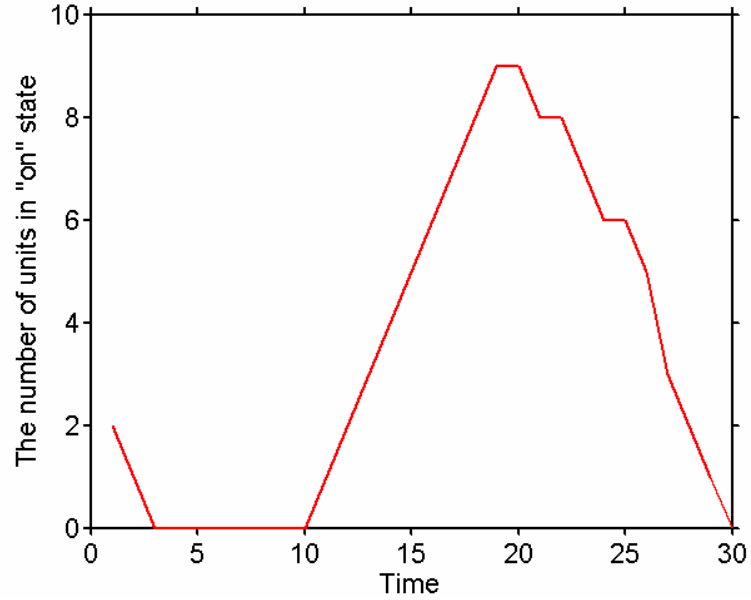
Figure 7.14. Number of Units in *On* State after a Price Increase

Three observations can be made based on these results:

1. There is a significant reduction of the load immediately after the lowering of the set points because the water heater units at a temperature higher than $T_+^{'}$ will turn off under the new setting. The remaining *on* units will turn off shortly, when they reach their new T_+ (Table 2.2, times 1, 2).

2. Because the cooling period is much longer than the heating period, the load output remains zero for a considerable length of time depending principally on how far the set point drops (Table 2.2, times 3 through- 8).

3. When the *off* units reach the lower temperature limit, they turn on, and the load begins to increase. Because of state degeneracy, there are more units in some states than in others'; the number of units that remain in "on" states is no longer a constant average value along the timeline. Assuming there is one unit in each state prior to the set point change, then after the set point change, the total number of "on" units will have a maximum at eight and a minimum at one, as shown in Table 2.2 and Figure 7.14 after time eight. The total number of *on* units is always five. Because $P_L=N_{on}P$ in the aggregated load profile, we expect a dynamic similar to that in Figure 7.14.

7.4.6 State Distribution in Response to Price Decrease

We can perform a similar analysis for the TCA response to a price decrease. We assume that the thermostat set point will be raised when the energy price drops. Raising the set point results in higher "off" temperature and "on" temperature limits of the TCAs. The degeneracy and the shift of the states are shown in Figure 7.15 and Table 2.3.

There are three observations based on the results:
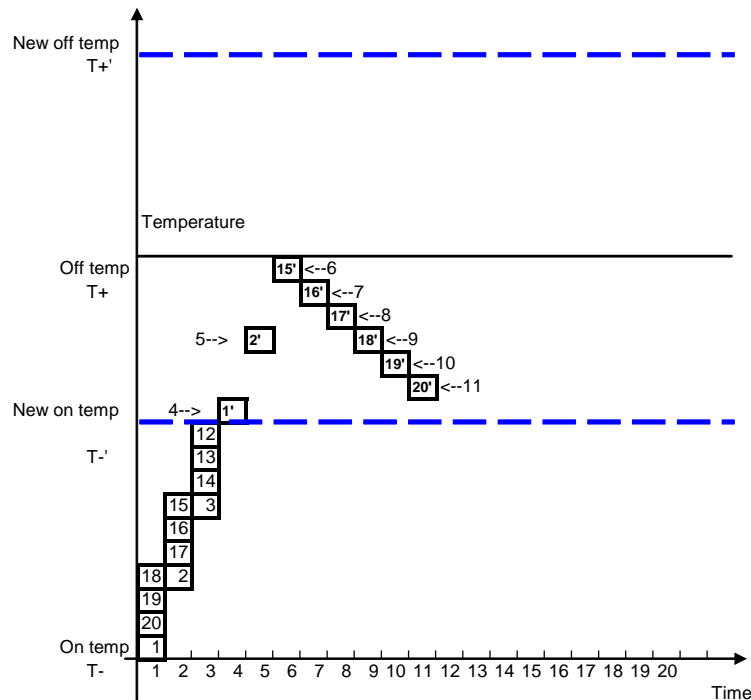
Figure 7.15.   State Degeneracy after a Price Reduction

1. Initially, there is a sharp increase of the load because all the TCAs at a temperature lower than $T_+^{'}$ will turn on under the new setting, while the units that were on remain on. If there are a large number of these units in a distribution zone, the peak load can reach much higher than the peak load that initiated the price increase.

2. The transient period is shorter than the previous case because the heating period is significantly shorter than the cooling period for water heater units.

3. Figure 7.16 shows that the synchronization behaviors are stronger than the previous case, which creates a much higher and longer peak and valley period in the overall load profile.

This reveals an interesting phenomenon. If the loads are price responsive and if we decrease the price prior to the expected peak price period, we may reduce the load during the peak price period *without the need for a price increase*. Pre-heating can be effective; however, the risk is that subsequently many water heater units will turn on and off at the same time. This can put a significant stress on the distribution network. In the next section, we will discuss this problem further.

7.14

Table 7.3.   State Distribution of TCAs after a Price Reduction

|	State	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	# of "on"
Transient State	1	on	on	on	1	2	15	16	17	18	19	20	1	on	on	on	on	on	on	on	on	14
	2	on	on	1	2	3	16	17	18	19	20	1	2	1	1	1	on	on	on	on	on	15
	3	on	1	2	3	4	17	18	19	20	1	2	3	2	2	2	1	1	1	on	on	16
	4	1	2	3	4	5	18	19	20	1	2	3	4	3	3	3	2	2	2	1	1	17
Steady State	5	2	3	4	5	6	19	20	1	2	3	4	5	4	4	4	3	3	3	2	2	17
	6	3	4	5	6	7	20	1	2	3	4	5	6	5	5	5	4	4	4	3	3	16
	7	4	5	6	7	8	1	2	3	4	5	6	7	6	6	6	5	5	5	4	4	12
	8	5	6	7	8	9	2	3	4	5	6	7	8	7	7	7	6	6	6	5	5	7
	9	6	7	8	9	10	3	4	5	6	7	8	9	8	8	8	7	7	7	6	6	3
	10	7	8	9	10	11	4	5	6	7	8	9	10	9	9	9	8	8	8	7	7	2
	11	8	9	10	11	12	5	6	7	8	9	10	11	10	10	10	9	9	9	8	8	1
	12	9	10	11	12	13	6	7	8	9	10	11	12	11	11	11	10	10	10	9	9	0
	13	10	11	12	13	14	7	8	9	10	11	12	13	12	12	12	11	11	11	10	10	0
	14	11	12	13	14	15	8	9	10	11	12	13	14	13	13	13	12	12	12	11	11	0
	15	12	13	14	15	16	9	10	11	12	13	14	15	14	14	14	13	13	13	12	12	0
	16	13	14	15	16	17	10	11	12	13	14	15	16	15	15	15	14	14	14	13	13	0
	17	14	15	16	17	18	11	12	13	14	15	16	17	16	16	16	15	15	15	14	14	0
	18	15	16	17	18	19	12	13	14	15	16	17	18	17	17	17	16	16	16	15	15	0
	19	16	17	18	19	20	13	14	15	16	17	18	19	18	18	18	17	17	17	16	16	0
	20	17	18	19	20	1	14	15	16	17	18	19	20	19	19	19	18	18	18	17	17	1
	21	18	19	20	1	2	15	16	17	18	19	20	1	20	20	20	19	19	19	18	18	3
	22	19	20	1	2	3	16	17	18	19	20	1	2	1	1	1	20	20	20	19	19	8
	23	20	1	2	3	4	17	18	19	20	1	2	3	2	2	2	1	1	1	20	20	13
	24	1	2	3	4	5	18	19	20	1	2	3	4	3	3	3	2	2	2	1	1	17
	25	2	3	4	5	6	19	20	1	2	3	4	5	4	4	4	3	3	3	2	2	17
	26	3	4	5	6	7	20	1	2	3	4	5	6	5	5	5	4	4	4	3	3	16
	27	4	5	6	7	8	1	2	3	4	5	6	7	6	6	6	5	5	5	4	4	12
	28	5	6	7	8	9	2	3	4	5	6	7	8	7	7	7	6	6	6	5	5	7
	29	6	7	8	9	10	3	4	5	6	7	8	9	8	8	8	7	7	7	6	6	3
	30	7	8	9	10	11	4	5	6	7	8	9	10	9	9	9	8	8	8	7	7	2
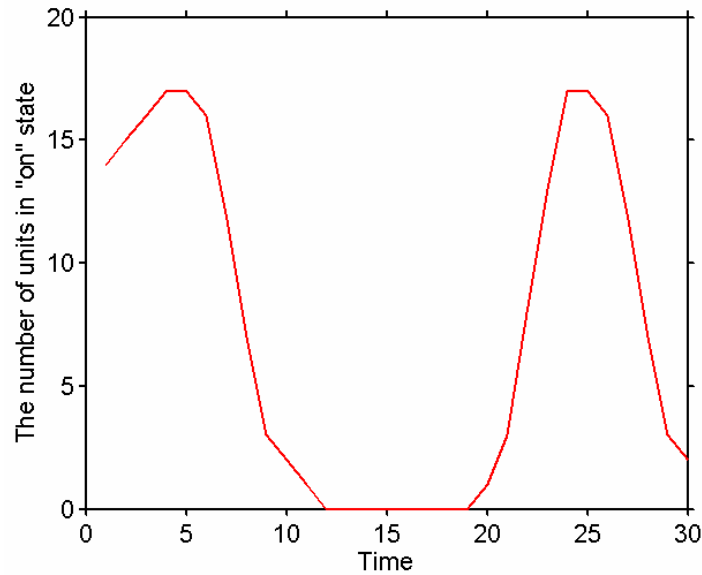
Figure 7.16.   Number of Units in *On* State after a Price Reduction

7.15

7.4.7   Price Response Impacts on Load Diversities

On a distribution feeder, the demand is not simply an aggregation of randomly distributed individual load demand; the correlation between loads must be accounted for. Diversity factors as a function of number of customers are used in distribution circuit loading calculations (Ghosh et al. 1997). Diversity factor (Johnson et al. 1997; Nazarko et al. 1998; Electricity Utility Engineering Reference Book, 1965) k_d is defined as

$$k_d = \frac{\sum_{i=1}^{n} D_i}{D_g} \tag{7.5}$$

where D_i is the maximum demand of load i and D_g is the coincident maximum demand of a group of n loads. k_d will have a value between $[1, +\infty]$. There are two extreme cases. If k_d is 1, all the individual loads are at *on* states. If k_d is infinity, all the individual loads are at *off* states. The lower the value k_d has, the more individual loads are at *on* states.

Assume that a feeder's load consists of N HVAC units. Each unit has a cycling time of τ with an *on* period of τ_{on} and an *off* period of τ_{off}. Also, assume that there are n states along the period of τ and each state evolves to the next state after the same deterministic time interval. Initially, if the units are uniformly distributed among each state, then the probability for a unit to reside in any of the n states is simply $1/n$. The diversity factor is calculated by

$$k_d = \frac{NP}{N \dfrac{n_{on}}{n} P} = \frac{NP}{N \dfrac{\tau_{on}}{\tau} P} = \frac{\tau}{\tau_{on}} \tag{7.6}$$

For the water heater example, there are 20 states consisting of 5 *on* states and 15 *off* states; therefore, the probability for a unit to be in any of the 20 states is 1/20 and the diversity factor is 4 (20/5).

To illustrate the impact of the price response on load diversity, we examine the case where TCA set points are lowered in response to a price increase. After lowering the TCA set points, two things will happen (see Figure 7.13). First, the states will immediately divide into in-regime states and out-of-regime states. Second, some states will be depopulated and some will have population increases. The set point changes will force the units whose temperatures are above the new off-temperature limit to turn-off; therefore they will "move" to the *off* states from the out-of-regime state. If the ratio of the set point change to the dead band is

$$k = \frac{T^+ - T^{+\prime}}{T^+ - T^-} \tag{7.7}$$

then the probability of finding a unit in out-of-regime states (States 6 through 14 in Figure 7.13) is

$$\Pr\{x > T^+\} = \frac{k n_{on} \dfrac{N}{n} + k n_{off} \dfrac{N}{n}}{k n_{off} N} = \frac{1}{n_{off}} \tag{7.8}$$
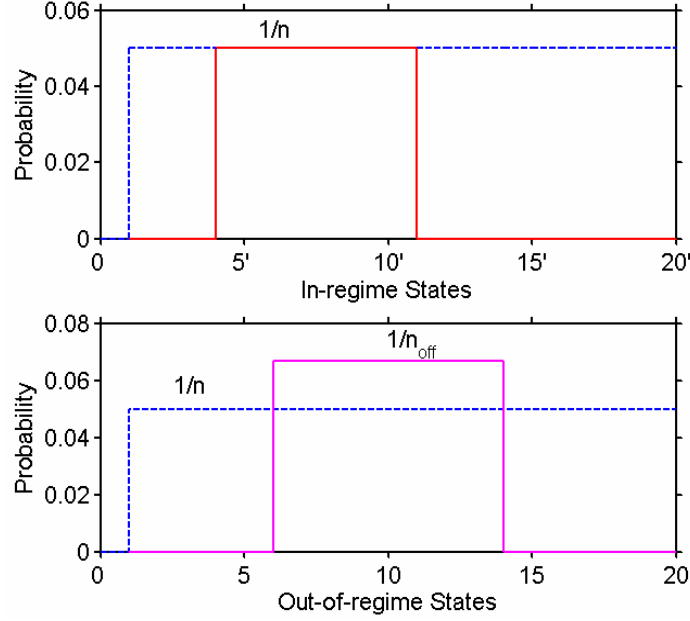
7.16

Figure 7.17.   PDF of In-Regime and Out-of-Regime States (for the Example described in Section 7.4.5)

where n_{off} is the number of *off* states (Figure 7.17). Note that $1/n_{off}$ is greater than $1/n$, showing an increased state occupancy for out-of-regime states. The remaining states with a temperature within the new settings are in-regime (States 4′ through 11′). Because the number of units in an in-regime state is unchanged, the probability for a unit to reside in an in-regime state is also unchanged and remains at $1/n$, as shown in Figure 7.17. However, those states below the old on-temperature limits are not occupied yet. So in a sense, they are depopulated.

The decay process will merge the out-of-regime states with the in-regime states and will repopulate some of the depopulated states among the in-regime states. The overall effect after the price response signal is that the PDF is no longer a uniform distribution but a superposition of the two PDFs.

The probability for a unit to be in state x after lowering TCA set points at the end of the transient is

$$
\begin{aligned}
&k\tau_{off} \le \tau_{on} - k\tau_{on} \\
&\Pr\{x \mid t \le k\tau\} = 0 \\
&\Pr\{x \mid k\tau < t \le \tau_{on}\} = \frac{1}{n} \\
&\Pr\{x \mid \tau_{on} < t \le \tau_{on} + k\tau_{off}\} = \frac{1}{n} + \frac{1}{n_{off}} \\
&\Pr\{x \mid \tau_{on} + k\tau_{off} < t \le \tau\} = \frac{1}{n}
\end{aligned}
\qquad
\begin{aligned}
&k\tau_{off} > \tau_{on} - k\tau_{on} \\
&\Pr\{x \mid t \le \tau_{on}\} = 0 \\
&\Pr\{x \mid \tau_{on} < t \le k\tau\} = \frac{1}{n_{off}} \\
&\Pr\{x \mid k\tau < t \le \tau_{on} + k\tau_{off}\} = \frac{1}{n} + \frac{1}{n_{off}} \\
&\Pr\{x \mid \tau_{on} + k\tau_{off} < t \le \tau\} = \frac{1}{n}
\end{aligned}
\qquad (7.9)
$$

where

$$
n_{off} = \frac{\tau_{off}}{\tau} n \ . \qquad (7.10)
$$

7.17

The probability of a unit staying in a specific state can then be calculated, and one can estimate the number of units in each state. By summing up the number of units in all the *on* states, the total power consumption can be estimated accordingly. However, because the units are no longer uniformly distributed in each state, the total number of *on* units is no longer a constant average value along the time line. The aggregated output will change with respect to time.

The whole process is more clearly represented by a queue with a deterministic service time, as shown in Figure 7.18 for the example given in Section 7.4.5. As shown in Figure 7.18a, assume that initially there is one unit (represented by a dot) in each of the 20 states (represented by a box). The set point changing process can be represented by two queues (Figure 7.18b): an out-of-regime queue and an in-regime queue. When the set point changes, previous States 1 and 2 become in-regime States 4 and 5, States 3 and 4 degenerate to out-of-regime States 9 and 12 and State 5 degenerate to in-regime State 6.

After all the units in the out-of-regime queue evolve to the in-regime queue, the transient process will be over (Figure 7.18c). The transient time is then $k\tau_{\text{off}}$. Based on the queue structures shown in Figure 7.18 and corresponding to different set point change ratio k in response to price increases, one can create an in-regime queue and an out-of-regime queue, as shown in Table 7.4, and simulate the whole system evolution along the timeline. The total power demand will then be determined by the number of units in *on* states at any time. Because the *on* state will always be in States 1 through 5, a sum of machines in these states will yield the total demand.

The inverse load diversity factor $1/k_d$ is calculated by dividing the aggregated load with maximum demand, *NP* (Figure 7.19). The diversity factor is no longer a constant as in the perfectly random distribution case. There are periods when the diversity factors are infinity, and we observe a synchronized *off* behavior. For some periods, the diversity factors are lower than the average at 4, where we observe stronger synchronized *on* behaviors. The results demonstrate that the price response has a significant impact on the load diversity. The remaining question then is whether and how the diversities are recovered after the response to a price change.

Table 7.4.   State Redistribution in Response to a Price Increase

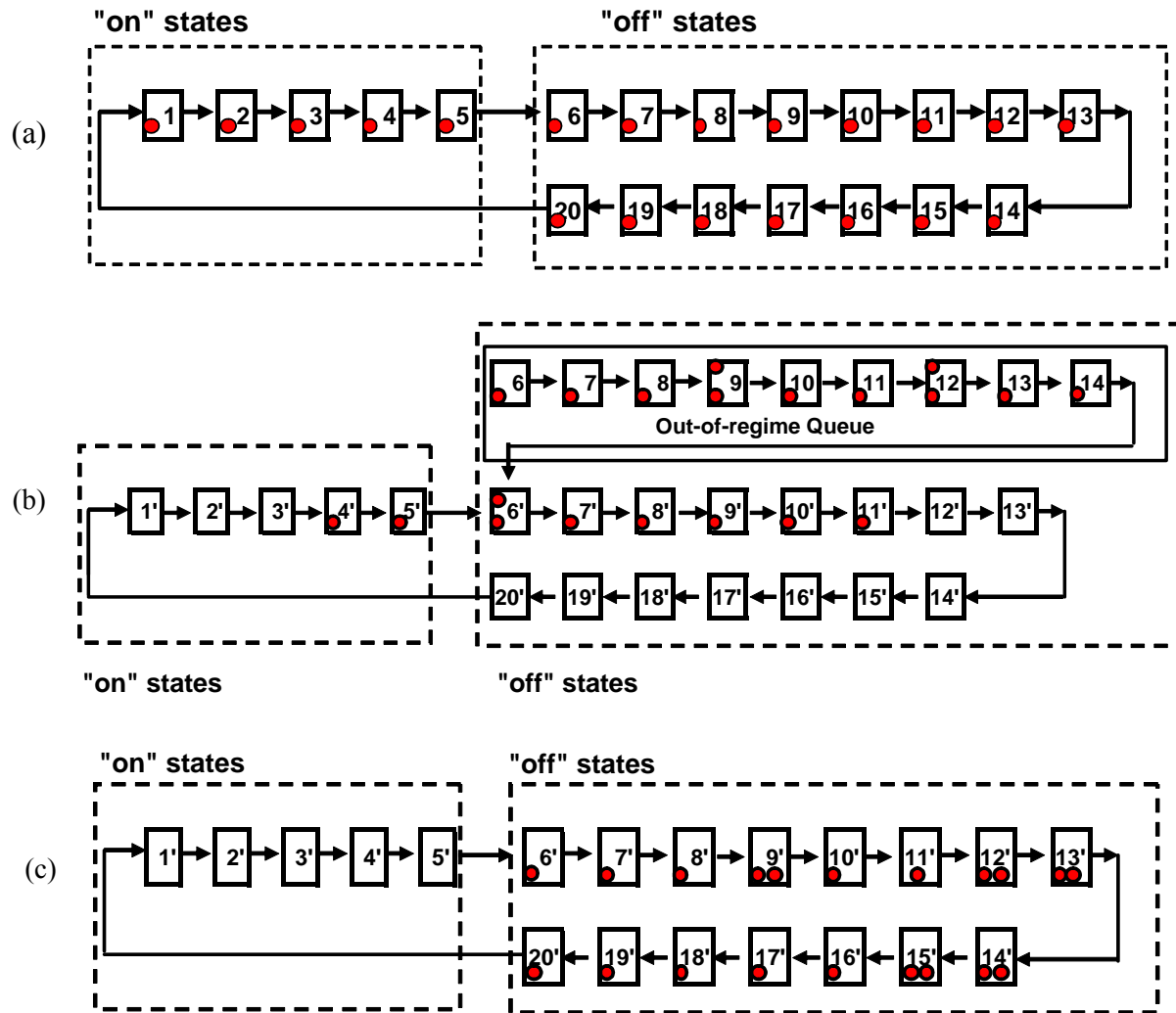| k	In-regime Queue	Out-of-regime Queue
0.2	2'-3'-5'-6'-7'-8'-9'-10'-11'-12'-13'-14'-15'-16'-17'-18'	6-7-8-6'
0.4	3'-4'-5'-6'-7'-8'-9'-10'-11'-12'-13'-14'-15'	6-7-8-9-10-11-6'
0.6	4'-5'-6'-7'-8'-9'-10'-11'-12'	6-7-8-9-10-11-12-13-14-6'
0.8	5'-6'-7'-8'-9'	6-7-8-9-10-11-12-13-14-15-16-17-6'
1	0	6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-6'
1.2	0	6-7-8-9-10-11-12-13-14-15-16-17-18-19-20-o-o-6'

Figure 7.18. A Queue Representation of State Transition of Price Increase Response Case
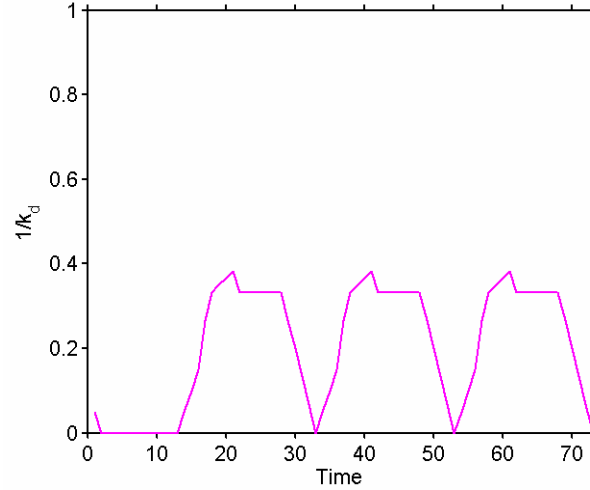
Figure 7.19. Diversity Factor after Price Increase

7.4.8 The Damping Process

Over time, random events such as the hot water usage or door openings of a refrigerator will naturally randomize any synchronous behavior of water tanks, HVAC units, or refrigerators and consequently damp the oscillation. Randomness brought by different types and sizes of TCAs as well as randomness caused by environmental differences may also contribute to the damping process. Our present model simulates the behaviors of hot water heaters having the same type and size with similar environmental conditions. Therefore, the damping in our model is mainly caused by random events. These random events cause state jumps of the unit in the queue. For example, taking a shower draws water from the hot water tank and so the tank temperature drops out-of-regime. For simplification, assume that each state except State 20 has a probability of p to return to State 1 and a probability of $1-p$ to evolve to the next stage. The queue can be modified as shown in Figure 7.20. Then, for a deterministic transition time, the transition matrix P is

$$
P = \begin{bmatrix}
p & 1-p & & & \cdots & \\
p & & 1-p & & & \\
\cdots & & & & \cdots & \\
p & & & & & 1-p \\
1 & & & & \cdots &
\end{bmatrix}
\tag{7.11}
$$

Therefore the state occupancy N after i transitions can be calculated as

$$
N_i = N_{i-1} P
\tag{7.12}
$$

When there are m on states, the number of units in on states will be

$$
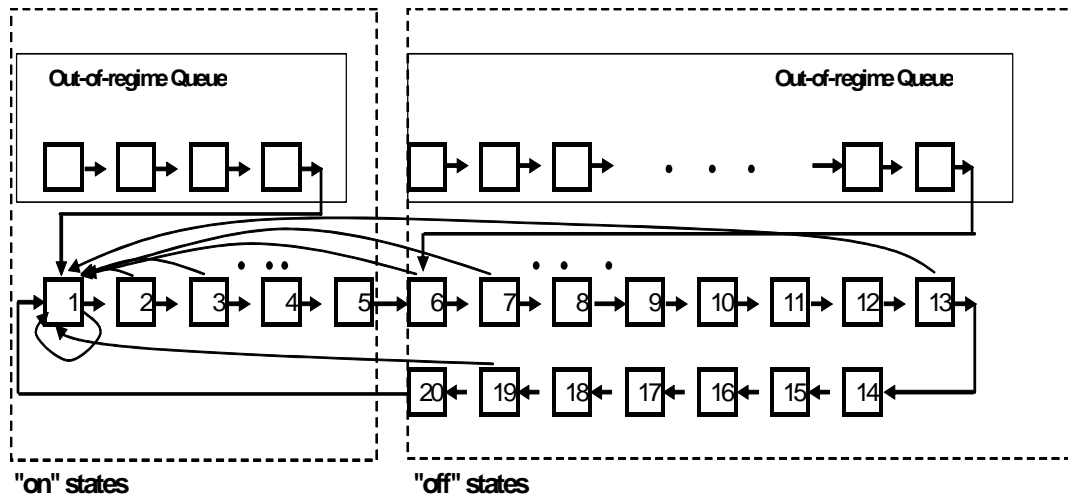N_{on} = \sum_{j=1}^{m} N_j
\tag{7.13}
$$

7.20

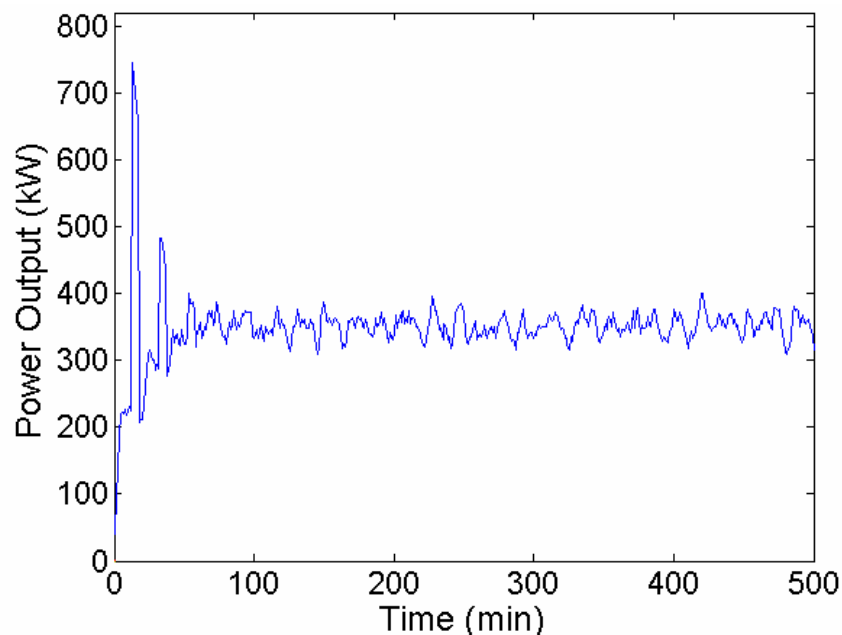**Figure 7.20.   Structure of the Queue Representation of TCAs**

Figure 7.21.   Damping of Load Response

Figure 7.21 shows the damping effect of the given example HVAC units on a feeder. The load dynamic is damped in about three cycles with a p of 1/20. The load diversity is then fully restored. In distribution systems, equipment are chosen based on an assumption that a load at the end of a feeder has certain load diversity. When all the loads turn on/off at the same time, the equipment maybe overloaded. Therefore, when designing a load response program, the natural damping factor of the system needs to be evaluated. If the system itself can not damp the dynamics fast enough, control strategies can be applied to diversify the load artificially.

7.5 Uncertainty

In this section we discuss the modeling of uncertainties in model parameters and randomness in customer consumption and their effect on the aggregate thermostatically controlled loads using a state queueing (SQ) model. The cycling times of TCAs vary with the type and size as well as ambient conditions. The random consumption of users, which shortens or prolongs a specific TCA cycling period, introduces another degree of uncertainty. By modifying the state transition matrix, these random factors were taken into account in a discrete SQ model. The impacts of considering load diversity in the SQ model while simulating TCA set point response were also studied.

To evaluate the economic benefits of different load shifting strategies, it is essential to model the change in TCA power consumption as a function of set point changes in response to market prices. A SQ model was presented in Section 7.4.3 to simulate aggregated TCA loads after a change in the set point. In Section 7.4.3, TCAs were assumed to be the same size and at the same set point under similar ambient temperatures, leading to a simplified model that only accounts for standby loss. However, in practice, TCA units are different sizes and the ambient temperature varies by region and season. Furthermore, random customer use has significant impact on the power consumption of TCAs. Using a water heater model as an example, methodologies are developed to modify the transition matrixes to account for the uncertainties in thermal parameters, the ambient temperatures, and the random customer behaviors, which are the key for the SQ model to simulate aggregated TCA loads accurately. The impacts of considering the randomness in the model on the TCA set point response are also studied.

7.5.1 Uncertainties in Thermal Model Parameters

The ETP approach is used to model the water heater consumption. An ETP representation of a water heater is shown in shown in Figure 7.22, where UA is the standby heat loss coefficient, C is the tank capacity, T_{out} is the ambient temperature, and Q is the heating rate. Q is proportional to the heater rated power. Based on the ETP model, starting from an initial temperature T_0 and ignoring the hot water consumption, the water temperature T can be calculated by

$$T = (T_0 - \frac{K_2}{K_1})e^{-K_1 t} + \frac{K_2}{K_1} \tag{7.14}$$

where K_1 and K_2 are calculated by

$$K_1 = \frac{UA}{C} \qquad K_2 = Q + T_{out}\frac{UA}{C} \tag{7.15}$$

To calculate τ_{on}, assuming $T_0 = T_-$ and $T = T_+$, we have

$$\tau_{on} = -\ln\left(\frac{T_+ - \frac{K_2}{K_1}}{T_- - \frac{K_2}{K_1}}\right)/K_1 = -\ln\left(\frac{T_+ - \frac{Q}{K_1} - T_{out}}{T_- - \frac{Q}{K_1} - T_{out}}\right)/K_1 \tag{7.16}$$

When the heater is turned off, Q is zero. The unit then coasts from T_+ to T. To calculate τ_{off}, we have
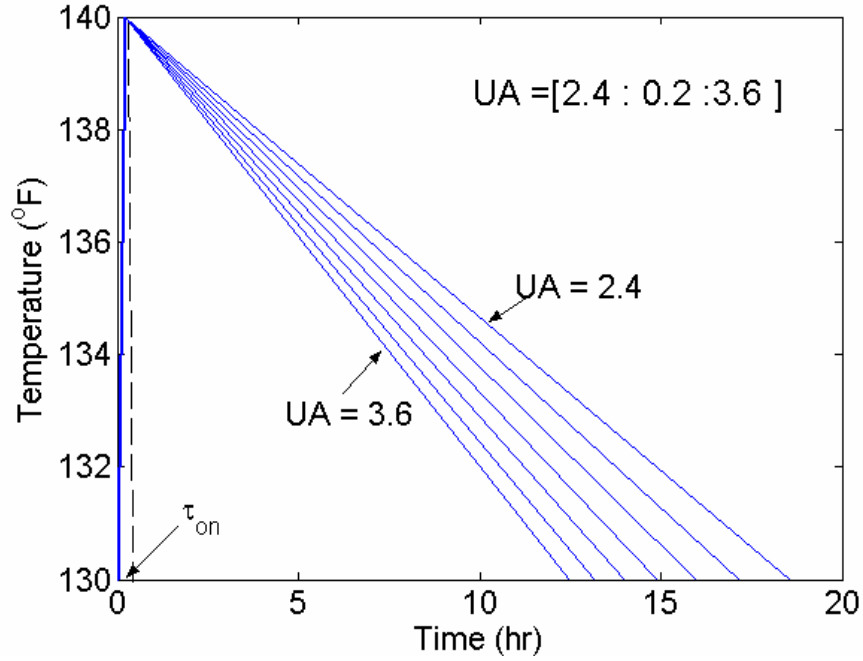
Figure 7.22. Thermal Characteristic Curves of Water Heaters with Different UA

$$Q = 0 \quad \Rightarrow K_2 = T_{out}\frac{UA}{C} \tag{7.17}$$

and

$$\tau_{off} = -\ln\left(\frac{T_- - \dfrac{K_2}{K_1}}{T_+ - \dfrac{K_2}{K_1}}\right) / K_1 = -\ln\left(\frac{T_- - T_{out}}{T_+ - T_{out}}\right) / K_1 \tag{7.18}$$

Based on (7.16) and (7.18), a sensitivity analysis can be performed to study the impact of parameter uncertainties on TCA cycling times. The analysis can be readily extended to the HVAC and the refrigerator models, which are similar to the water heater model.

7.5.2 Uncertainties in UA

The value of *UA* may vary even for the same type and size water heaters. To account for the variations, *UA* can be considered to follow a certain probability density function (PDF) in [UA_-, UA_+], such as a uniform or a normal distribution.

The heat rate of a water heater *Q* usually is much greater than the heat losses by design. Therefore, the value of *UA* will have a minor influence on the rising curve (τ_{on}), where the heater is on. When the heater is "off", *Q* is zero. Then, the heat loss determined by the *UA* of the water heater will dominate the length of the falling curve (τ_{off}).
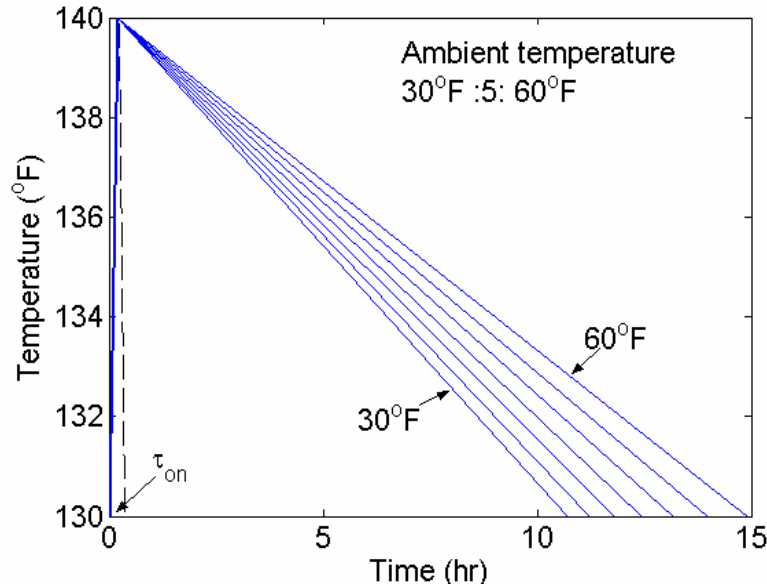
7.23

Figure 7.23. Thermal Characteristic Curves of Water Heaters as a Function of Ambient Temperature

As an illustration, consider 40 gallon water heaters rated at 4.5 kW with the *UA* varying from 2.4 to 3.6 Btu/hr°F. If the ambient temperature is 60°F, the calculated standby cycling periods are shown in Figure 7.22. As predicted, the τ_{on} values are almost the same for all units, but the τ_{off} values scatter from 12 hr to 18 hr.

At the feeder end, where the individual load aggregates, the differences in water heater cycling times caused by different *UA* of each unit result in an uncertainty in the aggregated thermal characteristic curve used to set up a SQ model. However, because *UA* is a deterministic parameter for each water heater unit, its impact is time invariant.

7.5.3   Uncertainties Caused by Different Ambient Temperature

Figure 7.23 shows an example of 40 gallon water heaters rated at 4.5 kW with *UA* at 3 Btu/hr°F. The ambient temperatures vary diurnally and seasonally from 30°F to 60°F. Note that this variation is a temporal variation, and its influence is a function of time. Again the τ_{on} values change slightly while the τ_{off} values scatter around a range of a few hours.

This is because for water heaters, the dead band is small compared with the value of *T-T_{out}*. Therefore, based on (7.18), τ_{on} will not change significantly with respect to T_{out}. Because the ambient temperatures vary with time, the uncertainties are time dependent.

7.5.4   Modified Thermal Model of TCAs

Taking into account the above uncertainties, the thermal characteristic curves for a group of TCAs are shown in Figure 7.24. The state that a unit resides in follows some probability distribution function determined by the uncertainties brought by the different UAs and ambient temperatures. When the ambient temperatures are high, there are more units having longer cycling periods and vice versa. The
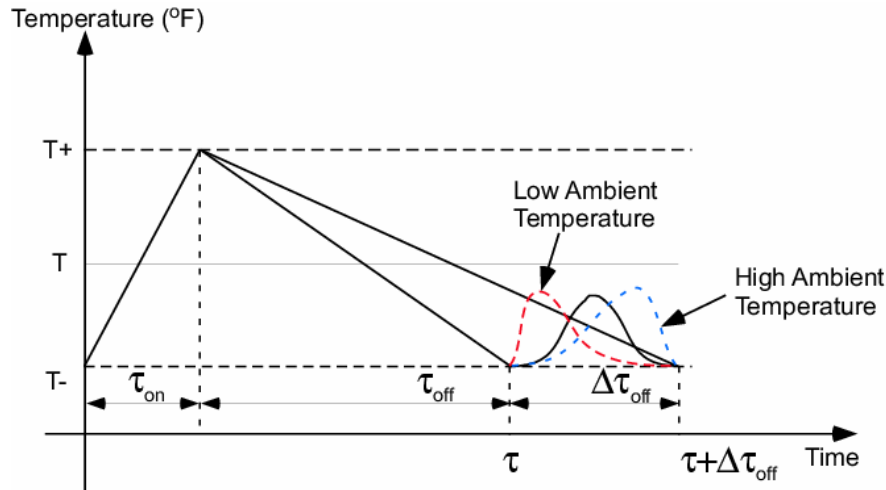
7.24

Figure 7.24. Uncertainties in Thermal Characteristic Curves

probability distribution will then be shifted accordingly, as shown in Figure 7.24. The uncertainties bring additional off-diagonal state transition probabilities in the probability transition matrix P, as shown in Figure 7.25. A SQ model considering the parameter uncertainties is shown in Figure 7.26, where N is the total number of states, and N_{on} is the number of *on* states.
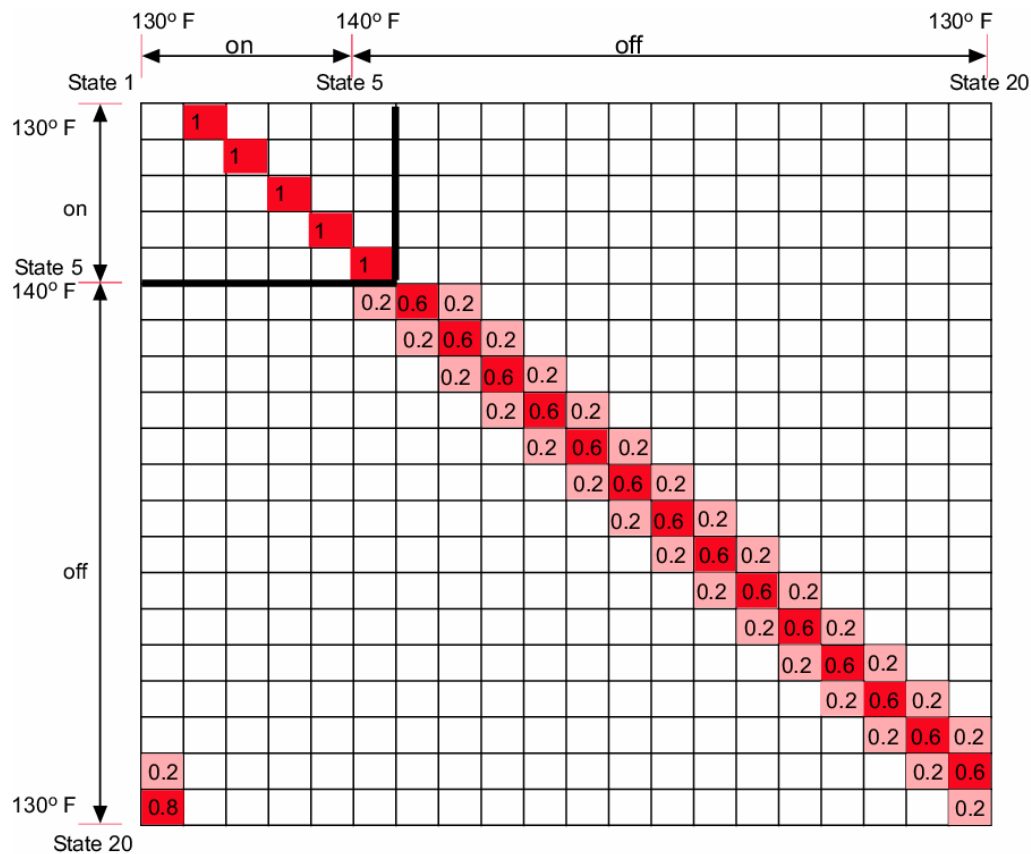

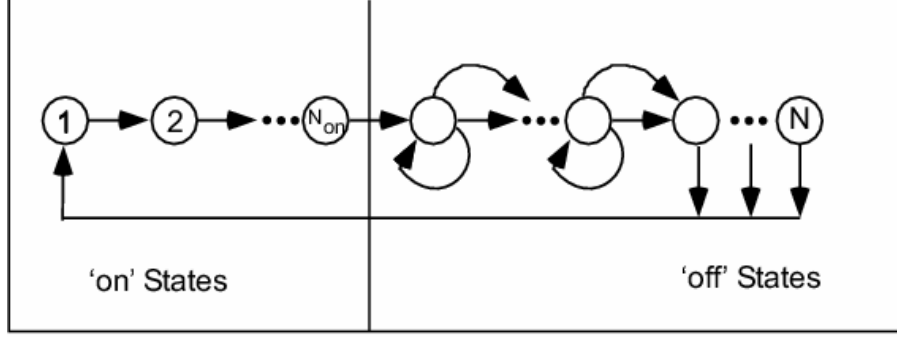**Figure 7.25. Uncertainties in State Transition Matrix P**

Figure 7.26. Modified SQ Model Considering the Uncertainties in TCA Cycling Time

Let vector $X_k = (x_1, x_2, .., x_N)$ represent the number of units in each state at the kth time interval and P represent the transition matrix (Gross and Harris 1998) containing the transition probabilities $p_{i,j}$, which represents the probability of a unit moving from State i to State j. We can then calculate the state evolution from the k-1 to the kth time interval using

$$X_k = X_{k-1}P \tag{7.19}$$

The aggregated power output P_L is then calculated by

$$P_L = P_{ave} \sum_{m=1}^{N_{on}} x_m \tag{7.20}$$

where P_{ave} is the average power output of a unit, x_m is the number of units in *on* state m, and N_{on} is the number of *on* states.

To account for the uncertainties, for units in state i, we have

$$\sum_{j=1}^{N} p_{i,j} = 1 \tag{7.21}$$

where $p_{i,j}$ follows certain PDF. For the 20-state transition probability matrix P shown in Figure 7.25, the unit residing in state 6 has probabilities 0.2, 0.6, and 0.2, to remain in state 6 or move forward to state 7 or 8. This shows that 20% of all units have slower or faster temperature decay rates than the majority, 60%.

In the next section, the impacts of customer random behaviors on aggregated TCA load curves are discussed.

7.5.5   Uncertainties Caused by Random Customer Behavior

Customer behavior can have a significant impact on the cycling time of TCAs. Based on load surveys, such as the End-Use Load and Consumer Assessment Program (ELCAP) (Peterson et al. 1993), PDFs of diurnal, weekly, and annual load profiles can be obtained.
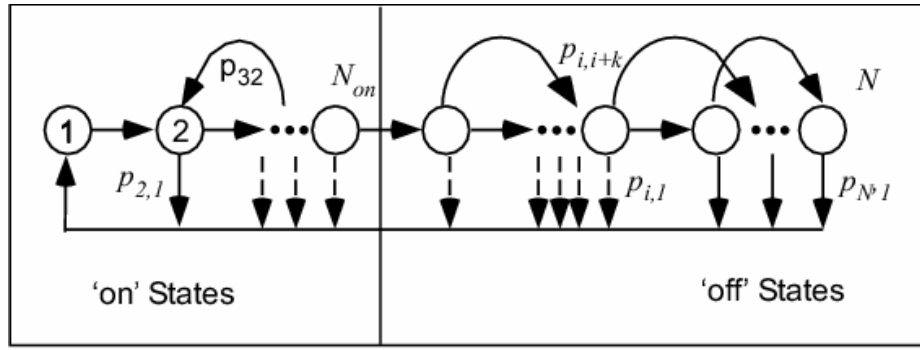
(a) The modified state queue
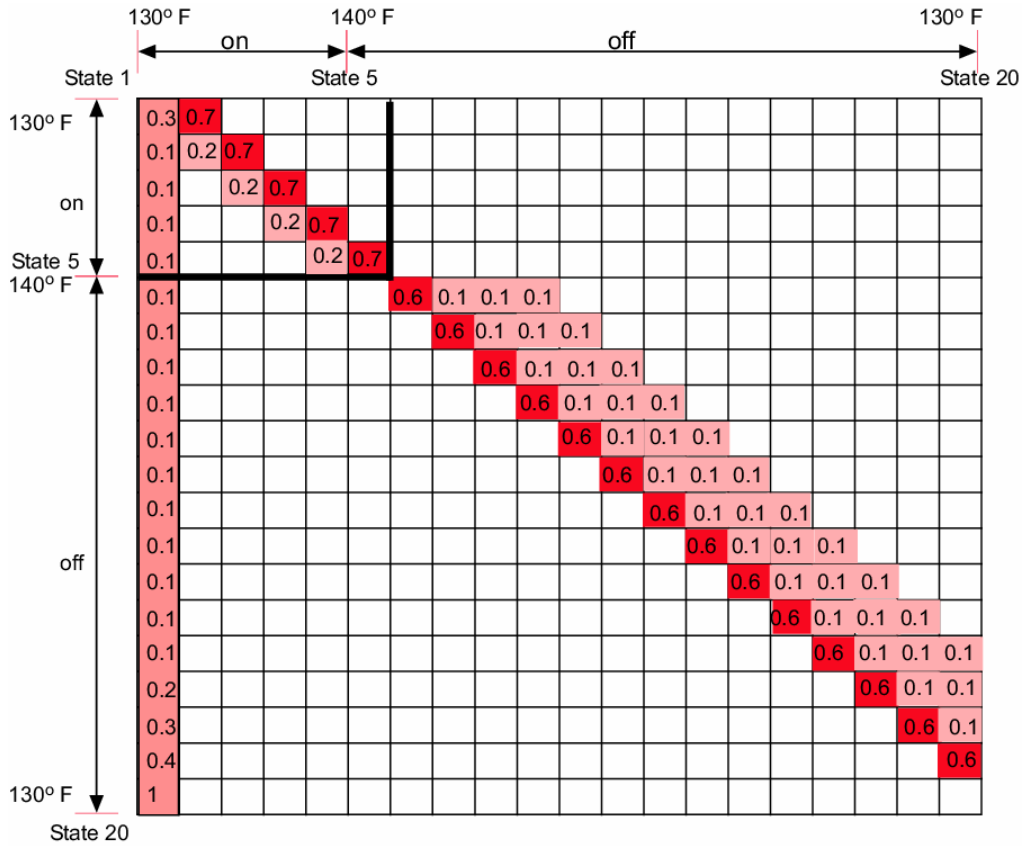
(b) The modified state transition matrix P

Figure 7.27. Modified SQ Model Considering Customer Consumption

Figure 7.27a shows an SQ model that takes customer hot water consumption into account. Minor consumption is a small amount of water draw, causing the water temperature to drop slightly. Because the temperature drops when the unit is on are not as significant as the temperature drops when the unit is off, the minor consumption can be modeled by adding $p_{i,i+m}$ ($m = 2, 3, \ldots$) entries for the off states and a $p_{i,i}$ entry for the on states in P (Figure 7.27b). Major consumption is a large amount of hot water draw that causes the water temperature to drop below the lower temperature limit. The heater will then turn on. This kind of consumption can be modeled by putting a link between each state to State 1, as shown in Figure 7.27b.
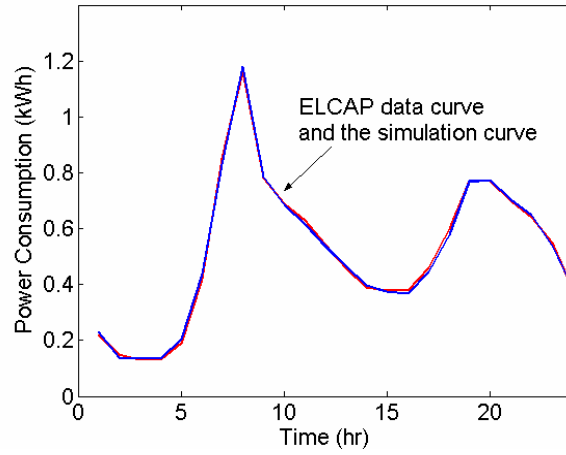
7.27

Figure 7.28. Comparison of ELCAP Winter Electric Water Heater Load Profile with Simulation Results

A transition matrix P corresponding to the SQ model shown in Figure 7.27b is

$$P=\begin{bmatrix} p_{1,1} & 1-p_{1,1} \\ p_{2,1} & p_{2,2} & 1-p_{2,1}-p_{2,2} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{N_{on},1} & & p_{N_{on},N_{on}} & 1-p_{N_{on},1}-p_{N_{on},N_{on}} \\ \cdots & & & \cdots & & \cdots \\ p_{i,1} & & & & 1-p_{i,1}-p_{i,i+2}-\cdots & p_{i,i+2} & p_{i,i+3} & \cdots \\ 1 & \end{bmatrix} \qquad (7.22)$$

where $p_{i,j}$ can be adjusted to account for the different consumption patterns. Note that $p_{i,j}$ is a function of time because the consumption of hot water varies by time-of-day. To tune the state transition probability $p_{i,j}$, we use the data collected in ELCAP by BPA in our simulation. The customer consumptions are classified into two categories: major consumptions and minor consumptions. A major consumption includes behaviors such as washing dishes, taking showers, and washing clothes, which usually last for more than 5 minutes. A minor consumption includes behaviors such as washing hands and washing fruit, which usually last for 1 or 2 minutes.

Figure 7.28 shows the ELCAP data curve versus the simulation data curve obtained using a 60-state SQ model with the probability curves shown in Figure 7.29. The probability of major and minor consumptions during a winter weekday is calibrated from ELCAP data. To tune the P matrix, we first estimate the probability of a minor consumption event based on consumption patterns. Then, we tune the major event probability to do the curve fitting. Because an eigenvalue analysis on the P matrix shows that the $p_{i,1}$ entries have a dominant impact on the steady-state distribution of the unit in each state for water heater case, the load profile follows the shape of the probability of major consumption events, as shown in Figures 7.28 and 7.29.
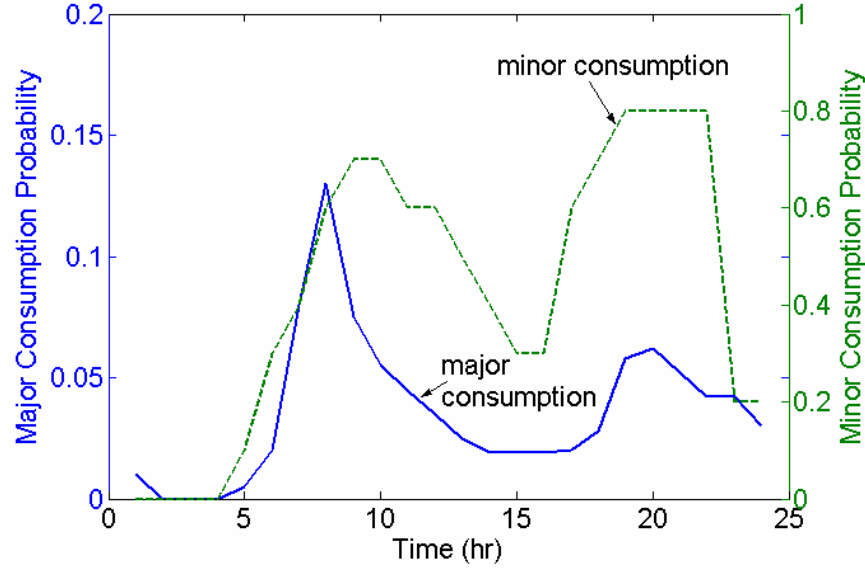
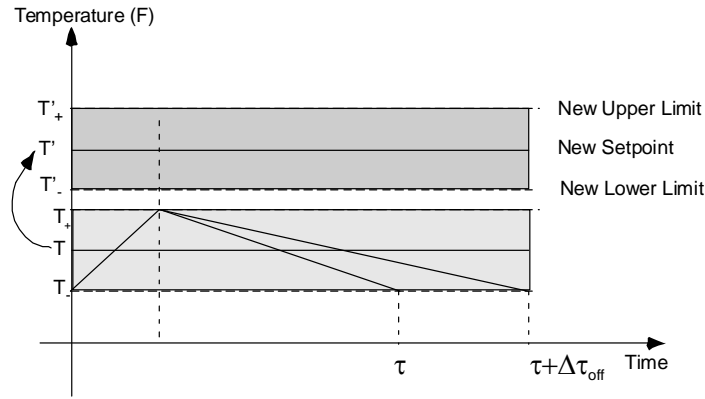**Figure 7.29.   Probabilities of Major and Minor Hot Water Consumption**

Figure 7.30.   Illustration of the Set Point Increase

7.5.6   Impacts on Set Point Change Response of TCAs

With a change in the set point, the upper temperature limit and the lower temperature limit of a TCA thermostat setting are shifted. Therefore, a unit may turn on/off accordingly, and electricity consumption is pre-consumed or deferred. However, the total energy consumed in 1 day remains the same.

7.5.7   Impact of Uncertainties in Load Cycling Times

Assume a fixed dead band during a set point change. When the market price drops, we expect a set point to increase, as shown in Figure 7.30, moving all units to below the new lower temperature limit. Therefore, the units that are off before the change of set point will turn on. Using a 20-state SQ model with 5 on states as an example and considering 10,000 units distributed uniformly along the 20 states, the average number of on units is 2,500. Considering only standby losses of the units, an oscillation on the aggregated load curve will be caused by a set point increase, as shown by the dotted line in Figure 7.31. The detail of this oscillation has been analyzed in Lu and Chassin (2004) using a TCA standby model.
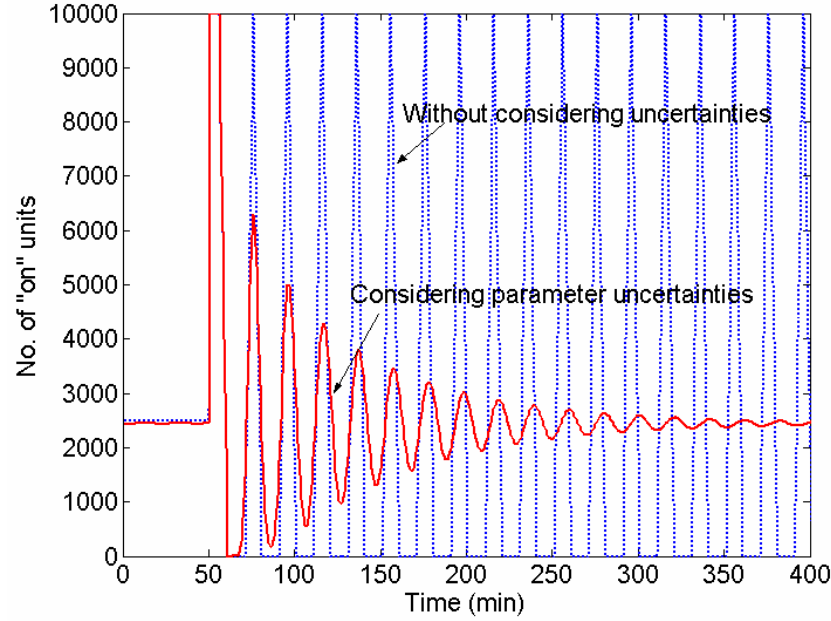
Figure 7.31. Number of Units in "ON" State after a Set Point Increase Considering Uncertainties in τ

In reality, the 10,000 units will not have exactly the same cycling time because of the differences in *UA*s and ambient temperatures. Using the modified transition matrix *P* shown in Figure 7.25 to account for these uncertainties, the aggregated TCA response curve then follows the solid line in Figure 7.31. We note following observations:

1. The initial peak is the same as that of the no-uncertainty case.

2. The succeeding peaks are damped.

3. The oscillation frequencies are different.

This is because the set point change basically will synchronize the units that are off with units that are on and form a dynamic oscillation with a frequency of $1/\tau$. The uncertainties in τ will cause a spectrum of oscillation frequency within $[1/\tau\ 1/(\tau+\tau_{\mathrm{off}})]$. Therefore, the aggregated load peaks will be the highest initially. After that, because of the frequency differences, the aggregated load peaks will be less than the initial one. The more diverse the unit cycling times are, the faster the aggregated load peaks are damped.

7.5.8 Impact of Uncertainties in Random Load Behaviors

The uncertainties in load cycling times are uncertainties in aggregated load behaviors. For each individual TCA unit, the cycling time is deterministic for a given ambient temperature. However, load consumption is different. Random demand may vary the cycling time of a TCA unit by randomly shortening the off cycle or by prolonging the on cycle, thereby increasing the effective duty cycle and bringing diversity to the water heater loads.
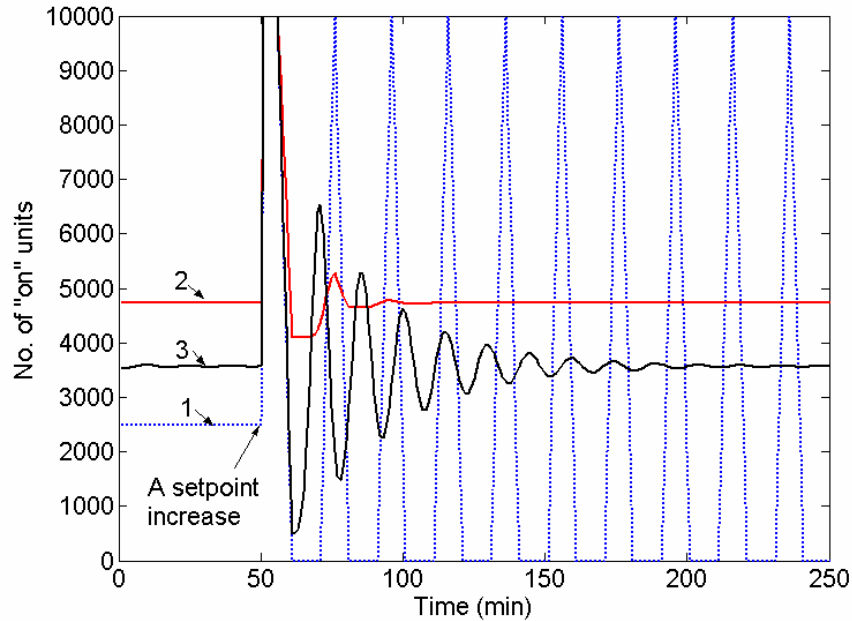
Figure 7.32. Number of Units in On State after a Set Point Increase Considering Customer Consumption

The 20-state water heater SQ model with five on states is again used for illustration. The probability matrix shown in Figure 7.27b is used to account for the customer consumptions. The set point increase response curve is shown in Figure 7.32.

As discussed in Section 7.5.5, demand varies with respect to time. During night hours, both major and minor customer consumption may be so infrequent that the probabilities of both can be considered zero. If a set point increase response is considered, we would expect the power consumption to follow curve 1 in Figure 7.32. In the morning hours, people frequently use hot water, and both minor and major consumption occur with high probabilities. If using the *P* matrix listed in Figure 7.27b, the response will follow curve 2 in Figure 7.32. During the day, when people go to work, the hot water consumption may be less frequent. Assuming that state transition probabilities for major consumptions change from 0.1 to 0.01 (as shown in Figure 7.27), the response will follow curve 3 in Figure 7.32. The set point decrease response for the three scenarios is shown in Figure 7.33.

There are several observations based on the results:

1. Demand raises the average power output because the effective duty cycle is shorter than the standby model.

2. Demand damps the load peaks after a set point increase and benefits the system by restoring the load diversity over time.

As shown in Figure 7.34, demand shortens the load shifting time obtained by lowering the set point. This may have a negative impact on load shifting programs because the load may not be able to be shifted away from the peak-price periods totally.
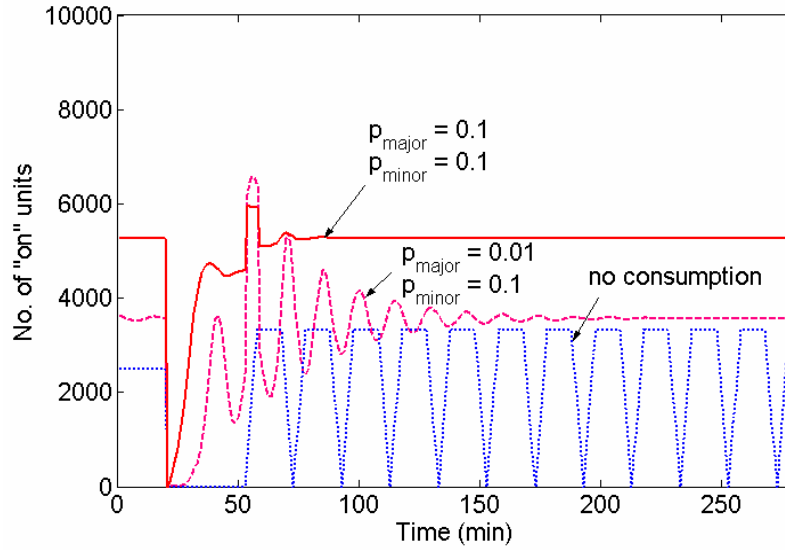
Figure 7.33.   Number of Units in On State after a Set Point Decrease Considering Customer Consumption
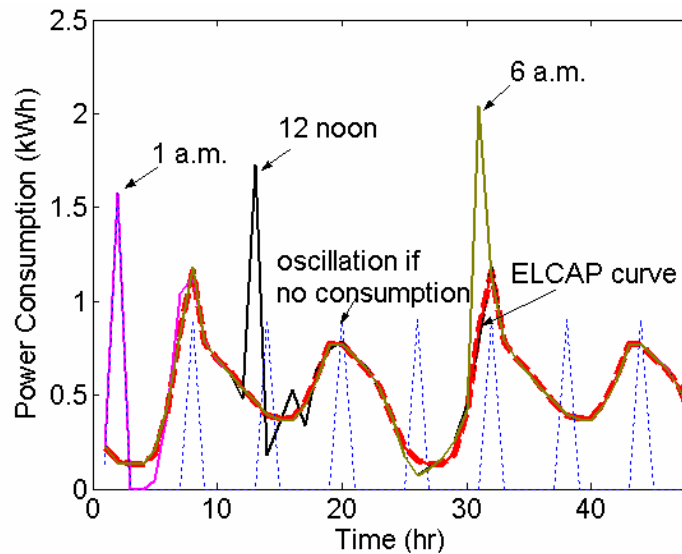

Figure 7.34.   Set Point Increase Responses Based on ELCAP Data

Figure 7.34 shows simulation results for the transition matrix tuned with ELCAP data. Assume that when prices drop, the set points of water heaters are raised. Then, consider a price drop at 1 a.m. The dotted line shows the load oscillations after this price drop if there is no customer consumption through-out the day. In real life, around 6 a.m., the increased hot water consumption starts to diffuse the syn-chronized loads and restore the original system diversity. If the price drop happens at 12 noon, when the minor consumption is high but the major consumption is low, then there is a small but observable oscillation in the load profile because of so much minor hot water consumption damp the oscillation. If set points are raised at 6 a.m., when both the major and minor hot water consumptions are frequent, we do not observe this oscillation because the damping rate is high. By calculating the eigenvalues of the transition matrix P, one can then find out how major consumptions can damp the load oscillations caused
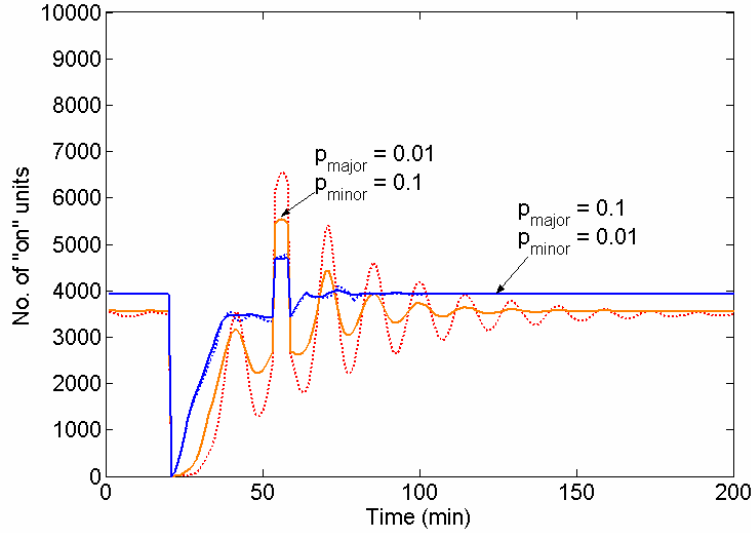
Figure 7.35. Number of Units in On State after a Set Point Decrease (combined uncertainty impacts)

by significant price drops or increases if the TCA set points are set to respond to price changes. The study can indicate whether distribution networks implementing such demand-side management strategies are overloaded under those circumstances and for how long.

7.5.9   Combined Impacts of Uncertainties in all Parameters

Figure 7.35 shows four cases of set point decrease responses. The solid lines are the responses considering both parameter uncertainties and the customer consumptions at different PDFs. The dotted lines show the cases considering only the customer consumptions. The results indicate that, although the randomness of the load cycling times and the customer consumptions both contribute to restore load diversities after a set point response, customer consumption has a dominant effect because higher customer consumption means faster diffusion processes between states. A sensitivity analysis on the eigenvalues and eigenvectors of the transition matrix P will also show that the higher the probability of customer consumptions, the higher the damping rates.

7.5.10   Comparison of Results from PDSS and SQ Model

The Power Distribution System Simulator (PDSS) (Guttromson et al. 2003) is a software simulation tool developed by Pacific Northwest National Laboratory (PNNL). PDSS is a simplified first-principles-based simulation tool that simulates individual loads with its own set of parameters. Then a load synthesis is done to get the aggregated load curve. Therefore, the uncertainties are accounted for at the unit level. The SQ model is both a state-based and an aggregated model. The uncertainties are accounted for at the system level. A comparison of results obtained using software developed under the two different approaches is depicted in Figure 7.36. The two software tools are first tuned to ELCAP hot water heater energy consumption data. Then they are used to simulate the load response to a set point decrease at 1 p.m. followed by a set point increase at 1 a.m. the next day. From the result, we notice that:
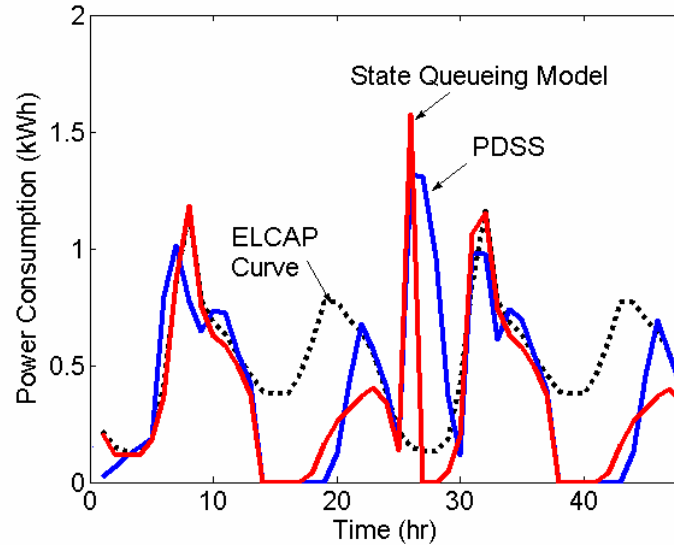
Figure 7.36.   Comparison of Results between PDSS and SQ Models

The responses are similar, although the magnitude and the phase may have some deviations. Because PDSS accounts for uncertainties at the unit level, the aggregated load obtained is more diversified than the SQ model. The synchronized load peaks obtained by PDSS are lower.

Compared with PDSS, the SQ model is easier to be tuned to follow a load curve exactly. This is because the input of the PDSS is at the unit level, and the aggregated load response at the system level is not fully controllable by the inputs.

When the population size becomes larger, for example, if a feeder with thousands of water heaters is simulated, the SQ model computes much faster because the computation time of a SQ model is sensitive to the number of states, but not to the total number of units.

8.0   Implementation

Demand-response (DR) programs increase demand elasticity, which can mitigate imbalance between supply and demand and moderate suppliers' ability to exercise market power to manipulate the price of electricity. In a competitive electricity market, there are generally two goals for the DR programs: mitigate price volatility and imbalance between supply and demand by load curve shaping. If the market price curve follows the load curve, the two goals lead to the same result, which is to consume power during low-price and light-load periods and to curtail or reduce power consumption during the peak-price and heavy-load periods. If the price peaks do not coincide with the load peaks, different control objective functions need to be applied. In this section, control strategies are developed to minimize the energy cost.

The strategies used in the DR programs can be classified into three different categories: curtailment, substitution (fuel switching), and load shifting (Bohn 1982; Schweppe et al. 1989; Daryanian et al. 1989). TCA control strategies used for load curtailment and shifting are discussed in this section. Load curtailment, as the name implies, curtails load during the anticipated peak-price or peak-load periods. Load curtailment cuts off customer or appliance electricity consumption during curtailment periods. It may or may not cause delayed consumption (additional consumption after the load comes out of curtailment). Load shifting, as the name implies, shifts electric usage to pre- or post-peak periods to reduce consumption during the anticipated peak-price periods. An important feature of the load shifting program is that it can target the cyclic loads such as TCAs. TCAs include residential HVAC systems, electric water heaters, and refrigerators. Varying the setting of a TCA thermostat can shift the TCA power consumption from tens of minutes to several hours, depending on the appliance. If the set point is controlled in response to the market price, the shifted TCA's power consumption contributes to load reduction for that period. In general, the TCA's power consumption is shifted rather than reduced; therefore, the electricity will have to be consumed either before or after curtailment period.

For evaluating the various control strategies, a sample comprising 1000 water heaters was chosen to represent a DR program implemented at a distribution feeder. The transmission level impacts of the DR programs can be studied by aggregating the feeder level load responses. The impact of different thermal characteristics of TCAs on their set point control strategies is first discussed. We then use electric water heaters as an example to evaluate the economic benefits obtained by the different control strategies. The disadvantage of each control strategy is also discussed in the paper.

8.1   Demand Response Programs in Electricity Markets

In a deregulated market, multiple parties in the bulk power systems engage in an open-access market competition with their own economic objectives. The market is bid-based and three time-sequential energy markets are established: the bilateral trade market, the day-ahead market (DA), and the real-time market (RT). The independent system operator (ISO) collects bids from generation companies (GenCos) and load servicing entities (LSEs), based on which a supply curve "*S*" and a demand curve "*D*" can be obtained. ISO then calculates the market clearing prices "*B*."

Figure 8.1 shows how an LSE can interact with the market by installing a price-responsive controller. Market clearing price *B* is a function of all *n* LSE bids, which can be represented as
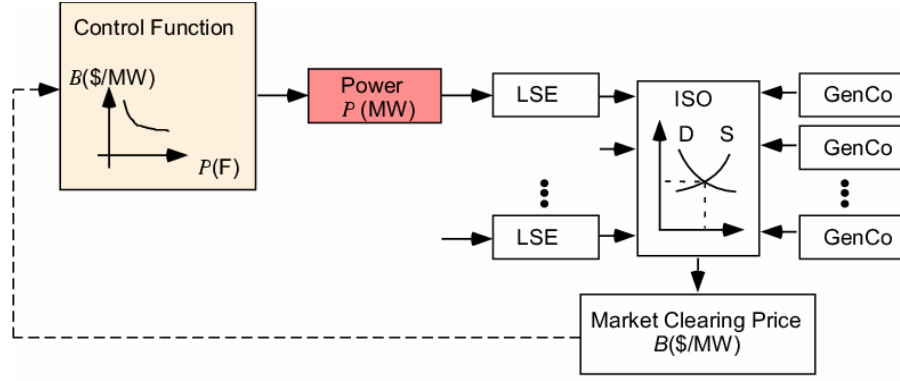
Figure 8.1.   Block Diagram of an LSE Bidding Process

$$B = f(P_1, P_2, ..., P_n)$$ (8.1)

where P_i is the supply bid of the ith LSE. The control function of the ith LSE is

$$P_i = f(B)$$ (8.2)

In principle, the control objective of an LSE is to maximize its profit over a time period T. However, it is typical for today's LSE to consider load as a given, and thus they seek only to minimize their costs. The optimization problem is then reduced to

$$\min\left(\sum_{t=0}^{T} P_i(t)B(t)\right)$$ (8.3)

To determine the load response P_i at a specific time t, two cases need to be studied. If the market penetration of the DR is low, an LSE will not have control over B. The market is then considered to be a competitive power market, where the $B(t)$ is insensitive to the power variation of a single load bidder. In this case, there are no feedback loops between the market price B and the load power P_i in the simulation. The dotted line in Figure 8.1 is then disconnected. At each time step, the LSE will calculate the price B and the load power P_i iteratively based on Eq. (8.1) and (8.2) to meet the control objective (8.3, which minimizes the load payment at t.

When the DR programs have a higher penetration, so that the load reduction or increase is significant at the transmission level, B will be influenced by an individual bid P_i. In this case, the dotted line in Figure 8.1 is then connected. At each time step, the LSE will calculate the price B and the power P_i iteratively based on Eq. (8.1 and (8.2) to meet the control objective (8.3), minimizing the cost of energy at t.

If distribution generation is considered in the cost optimization by contracting directly with the LSEs, the optimization problem can be reformulated as
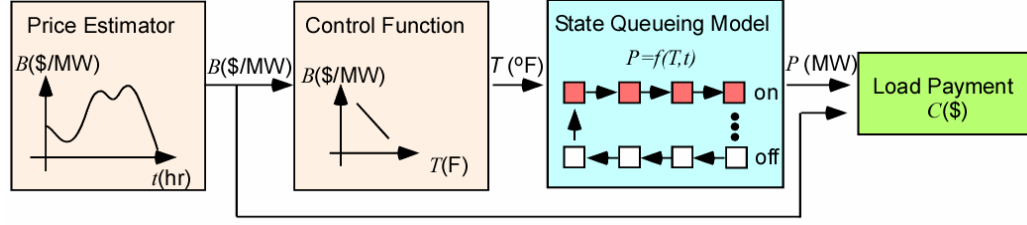
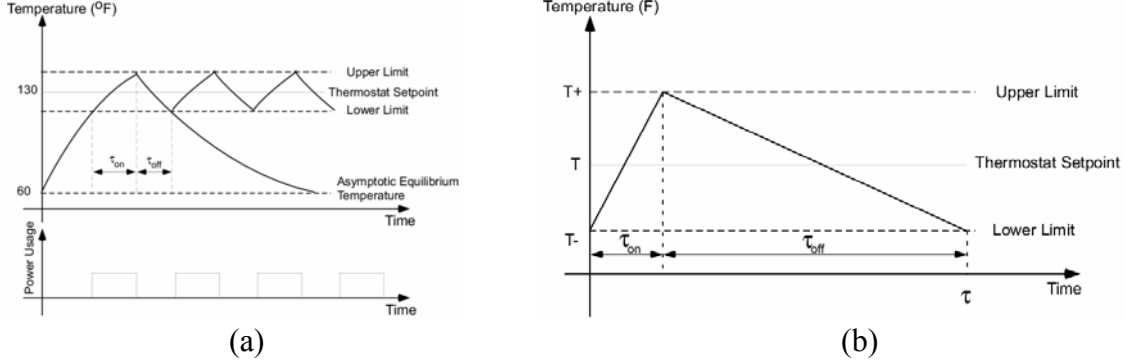**Figure 8.2.   Block Diagram of Load Payment Calculation**

(a)   (b)

Figure 8.3.   Thermal Characteristic Curve of a Water Heater

$$\min(\sum_{t=0}^{T}(P_i^w(t)B_w(t) + P_i^c(t)B_c(t)) \quad \text{(8.4)}$$

where B_w is the wholesale market price, B_c is the contract price with the distribution generators, P_i^w is the power bid into the wholesale market, and P_i^c is the power contracted with the distributed generators.

From historical data, an LSE can estimate a daily DA market price curve. Then, based on the control function assigned to TCA thermostats, the aggregated TCA power consumption can be obtained and the load payment calculated. An economically optimized control strategy will yield minimum load payment with all the physical constraints satisfied.

Physical constraints exist because electrical distribution components such as cables, transformers, and fuses are rated according to their capacity to transmit power. For feeder circuits serving predominantly residential loads, selection of equipment is often based on the peak diversified load (Lang et al. 1982). Thus, the maximum aggregated load on a feeder puts a constraint on the optimization. The whole LSE cost optimization process is shown in Figure 8.2.

8.2   Thermal Models of TCAs

As stated earlier, Figure 8.3a shows the thermal behavior (temperature of stored water) of a water heater unit over time. The rising curves indicate the water heater is on, and the falling curves represent the standby (or cooling down) periods, when the heater is off. As the water heater unit cycles, the water temperature in the tank rises and falls accordingly. The upper and lower limits represent the deadband of
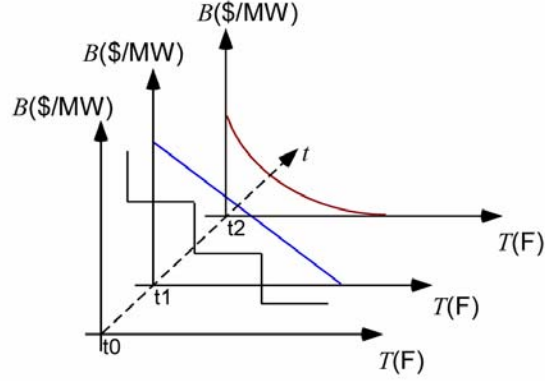
Figure 8.4.   Various Thermostat Set Point Control Functions

the thermostat around the thermostat set point, and changing the set point allows one to regulate the power consumption of the TCAs. Because the asymptotic equilibrium temperatures are generally far beyond these limits for appropriately sized equipment, the exponential rising and falling curves are nearly linear between the upper and lower limits. To simplify the analysis, a linear approximation of TCA thermal characteristics (Figure 8.3b) is used in our model. T_+ and T_- are the upper and lower temperature limits for a given set point, T. The unit has a cycling time of τ, with an on period of τ_{on} and an off period of τ_{off}.

The thermostat setting T is the preferred set point of the customer. Different TCAs have different ranges for their thermostat settings, which reflect customer's comfort choice. For water heaters, the thermostat settings vary from 120° to 160°F (HowStuffWorks 2004). However, settings above 140°F may cause scalding if the water is not mixed with cold water prior to use. The recommended temperature range for food is between 34° and 40°F and for frozen food between 0° and 5°F. Otherwise, there is a greater chance of rapid food spoilage and bacteria growth (Bauer et al. 2001). A customer may choose to temporarily reset the set point and make a sacrifice in comfort to reduce his energy bill. Nevertheless, at certain times the customers are willing to pay whatever the price is to maintain the comfort level. Therefore, the thermostat settings $T=f(b,t)$ is modeled as a function of time as shown in Figure 8.4, where B is the market price and t is time.

The wider the set point variation range, the longer period the load can be shifted. The cycling times for refrigerators are around one hour, while for air-conditioners and heat pumps they can be as high as one or two hours,; and for water heaters it can be up to twelve hours. Therefore, a water heater is a good candidate for DR programs. Another reason is that water heater load can be viewed as non-weather-sensitive loads (Gellings and Taylor 1981) because the set point temperature is far above the equilibrium temperature.

8.3   Evaluation of Control Strategies

If the goal of the LSE is to implement a DR program for the water heaters that would minimize cost of energy in DA market, then there are two control strategies: load curtailment during the peak-price period and preheating to coast through the peak-price period.
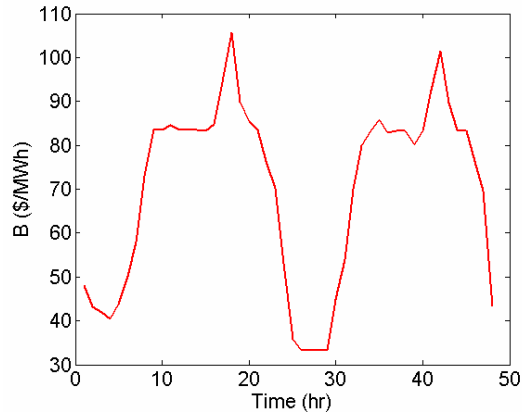
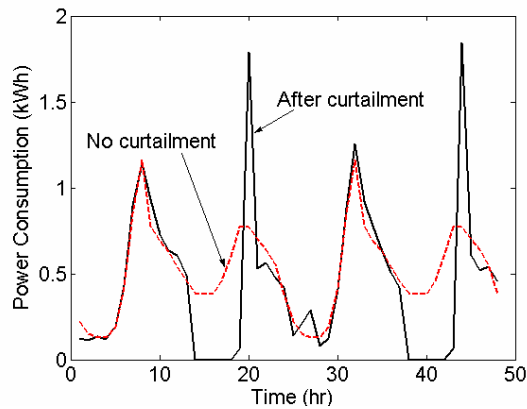**Figure 8.5.   Day-Ahead Market Price Curve for Two Days**

Figure 8.6.   Water Heater Unit Power Consumption

8.3.1   Load Curtailment

Figure 8.5 shows a DA market price curve obtained from NYISO.[a] Based on historical data, an LSE can estimate DA market price curve and target the curtailment in peak-price hours. Figure 8.6 shows the hourly load profile when implementing load curtailment from 1 p.m. to 7 p.m. on a group of water heater load. The rated power of an individual water heater is 4.5 kW, and its thermostat setting is 120°F. The results are normalized to a single water heater load.

There are several observations based on the simulation results:

1. There is a payback period[b] where the power consumption rises sharply. This is because when power is restored, water heaters will turn on simultaneously and the load diversity has been lost. The power surge can put significant stress on distribution circuits as pointed out in (Lee and Wilkins 1983).

(a) http://www.nyiso.com
(b) Payback period is defined as the rebound of power consumption that exceeds the normal consumption because the state of the systems is below what is normal.
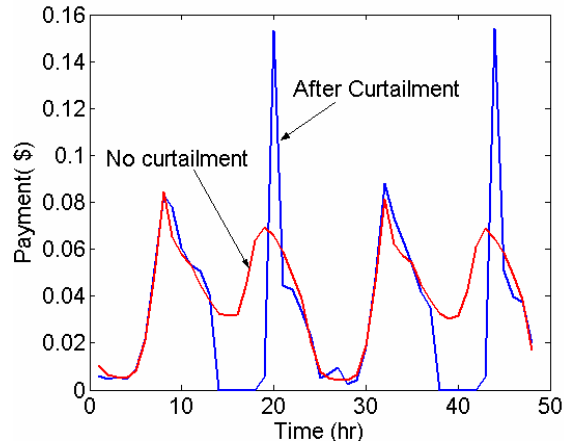
Figure 8.7.   Water Heater Unit Payment

2. As load curtailment cuts off the power supply completely for hours, customer comfort was sacrificed in exchange of the economic benefit.

3. After the curtailment period ends, between 8:00 p.m. and 9:00 p.m., customer consumptions are fairly active, which helps the system to restore load diversity.

Figure 8.7 shows a comparison of energy cost with and without load curtailment. As long as the daily total cost of energy for the curtailment case is less the total daily cost of energy during normal operation, the program is economically beneficial to the customer and the LSE.

8.3.2   Preheating and Coasting

Because load curtailment will limit customer energy consumption during peak-periods and may cause inconvenience, wide spread adoption may be limited. Preheating the water to a higher temperature than normal by increasing the set point temperature, one can receive the same power reduction.

As reported earlier, the thermostat settings of water heater varies between 120° and 160°F. One can set the set point to a higher one and then lower it to reduce the power consumption during the peak-price period. We first implement a very simple control strategy, which is to set the thermostat at 140°F at 11 p.m. and then lower it to 120°F at 1 p.m. everyday.

Figure 8.8 shows the set point-control does achieve the same power reduction over the targeted time period. The rising curves are also smooth. However, in order to get ready to preheat the units, the set points have to be raised to 140°F at 11 p.m. This set point increase brings a power surge similar to the payback phenomena after a load curtailment (Figure 8.9). However, as the market price is lower for mid-night hours, the cost impact is not as high as that of the load curtailment case (Figure 8.10).

With set point-control, comfort is not sacrificed because the set point during peak-period is un-changed. If the water temperature drops below the lower temperature settings, the heater will turn on. In Figures 8.9 and 8.10, we can see that the power consumption is not zero during the peak-period. This is because hot water consumption will cause the water heater to turn on for short periods.
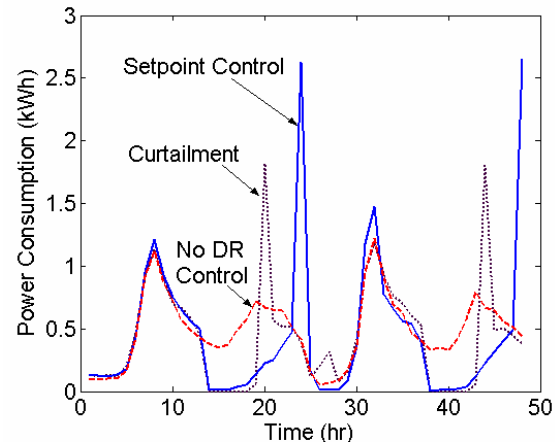
8.6

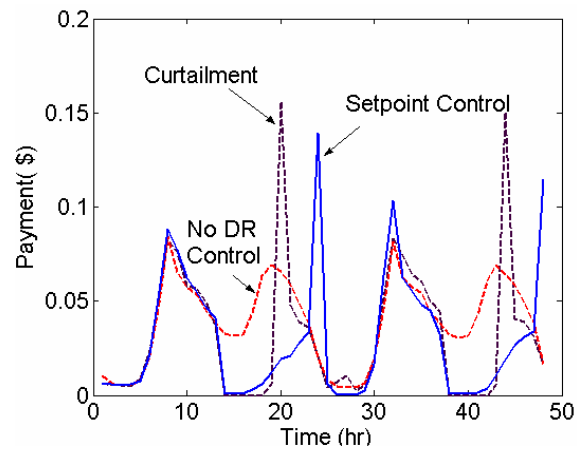**Figure 8.8.   Energy Consumption of a Water Heater**

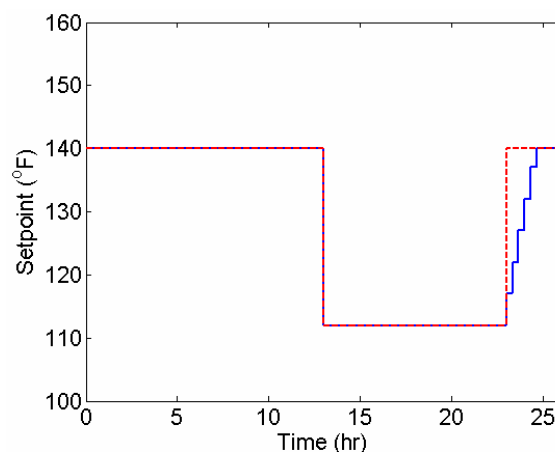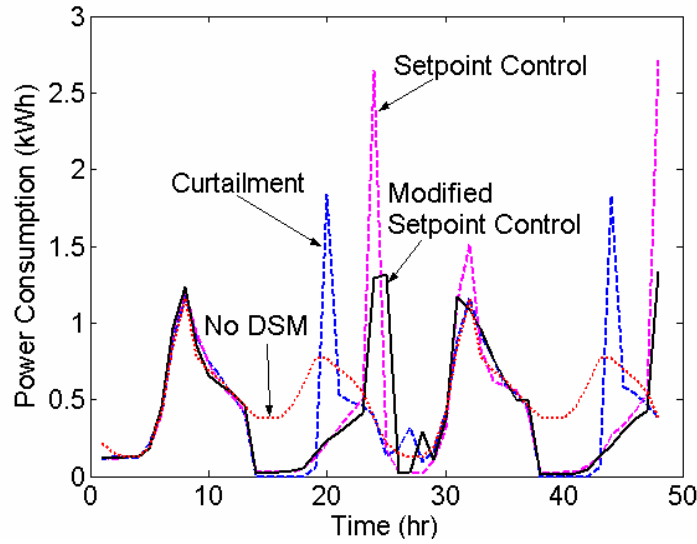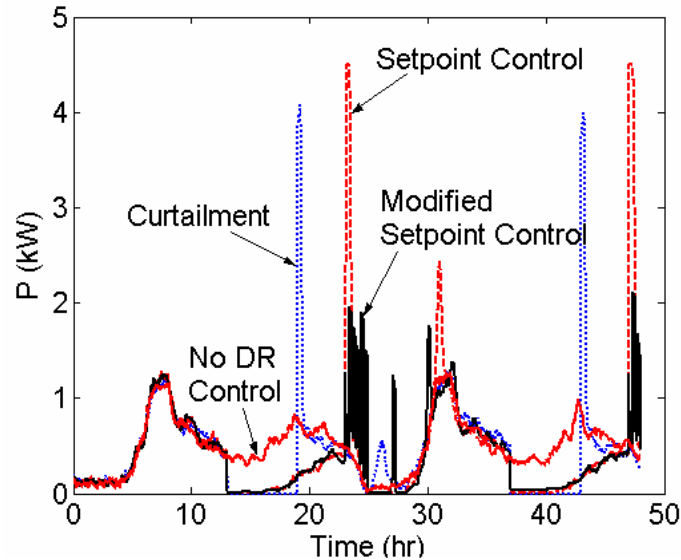**Figure 8.9.   Energy Payments of a Water Heater**

Figure 8.10.   A Modified Set Point-Control Curve

(a) The hourly power consumptions of a water heater

(b) The power consumptions of a water heater using 1-minute step ramp

Figure 8.11.   Water Heater Energy Consumptions under Different Control Strategies

    The problem then is how to control the power surge at the beginning of the preheating period. There are two ways to raise the set point: randomizing the turn-on sequence and a step by step increase of all the thermostats. Randomizing the turn-on sequence needs additional communication equipment between the water heater unit and the controller (Weller 1988). Another way is to increase the set point in small steps, for example, 5°F every 20 minutes, as shown in Figure 8.10. The water heater power consumption and energy cost are shown in and Figure 8.11 and 8.12. In Figure 8.11a, we see a significant reduction in peaks after the set point is raised with modified set point approach. This is because the set point increase has been broken into several steps. At the beginning of each preheating stage, there is a smaller power, as shown in Figure 8.11b. As the equipment in distribution network is rated to meet the peak diversified load, this modified control strategy of peak load reduction will greatly relieve the stress of overloading cables and transformers and avoid tripping the relays.
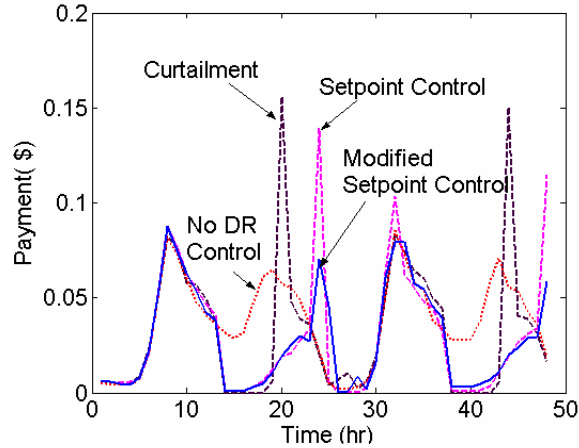
Figure 8.12. Water Heater Energy Payments under Different Control Strategies

Table 8.1 shows the hourly energy reduction achieved by applying different water heater control strategies, and Table 8.2 shows the daily power consumption E and the daily energy cost C of a water heater under different control strategies. The results indicate:

1. Although the energy reduction of set point control is less than the curtailment, it does allow the customers to use their appliance during the control period. Therefore, the comfort of the customer is not compromised.

2. Because the set point-control strategy can extend the payback period to hours when the market price is low, the overall energy cost is less than that of the load curtailment strategy.

3. Notice that the savings are in the order of tens of cents per day, meaning a possible few dollars per months saving to a customer. This indicates that TCA control strategies have to be inexpensive to implement.

Table 8.1. Water Heater Hourly Energy Consumption During Curtailment

Control Strategies	1 p.m.	2 p.m.	3 p.m.	4 p.m.	5 p.m.	6 p.m.	7 p.m.
No Control (kW)	0.43	0.38	0.35	0.37	0.49	0.57	0.72
Curtail (kW)	0.45	0.00	0.00	0.00	0.00	0.00	0.06
Setpoint Control (kW)	0.49	0.01	0.01	0.01	0.03	0.05	0.14
Modified Setpoint Control (kW)	0.46	0.01	0.02	0.01	0.03	0.05	0.11

Table 8.2.   Water Heater Daily Energy Consumption and Energy Cost

Control Strategies	*E*(kWh)	*C* ($)
No control	12.75	1.02
Curtailment	11.22	0.85
Setpoint Control	11.51	0.83
Modified Setpoint Control	11.24	0.80

8.4 Optimal Control Strategies

The preheating and coasting control of a TCA unit is similar to the control of a pumped-storage power station. The preheating period is a heat storage process, which is similar to the pumping mode of a pumped-storage power station and the coasting period is similar to a water discharging mode of it. Therefore, optimal bidding strategies developed by Lu et al. (2004e) for a pumped-storage hydro-turbine can be used for LSEs applying the preheating and coasting control over their TCA units.

In order to minimize the payment, the preheating period should be done during the least price hours and the coasting period should cover the peak-price period. Because the coasting time is in a range of a few hours, the optimization will be done on a daily basis. Therefore the optimization process is as follows:

Step 1: Obtain an estimated day-ahead market clearing price (MCP) curve.

Step 2: Target the coasting period to the hours when the prices are the highest and target the preheating period to the hours when the prices are the lowest. Based on the length of the coasting period, calculate the range of the set point need to be adjusted. Calculate the power output under such a set point regulation. Calculate the total payment *C*.

$$C = \sum_{i=1}^{24} B(i)P(i) \tag{8.5}$$

where $B(i)$ is the MCP for the ith hour and $P(i)$ is the total power output at the ith hour.

Step 3: Extending the coasting period to the next highest price hour and calculate the total payment. If the coasting time reaches the maximum TCA load shifting time, stop. The maximum TCA load shifting time is determined mainly by the range of the TCA set point.

For the above example, if the water heater set point can be regulated between 120° and 140°F, the coasting time is around 10 hours. With the estimated MCP curve shown in Figure 8.10, the optimal control strategy will be preheating the water heaters at 2 a.m. by setting the set points to be 140°F and letting them coast after 10 a.m. by dropping the set points to 120°F. The resulting load profile and load payment is shown in Figures 8.13a and b. The dash line indicates the load profile and the energy cost without implementing TCA set point-control.

The results indicate that the greater the price differentials are in a day, the more savings the TCA load control can bring to an LSE.
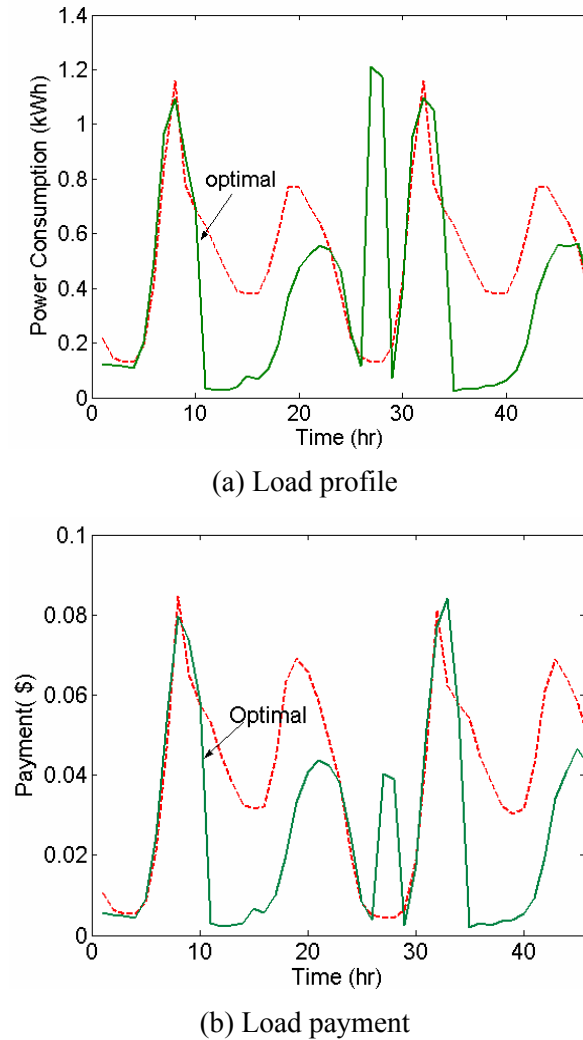
(a) Load profile

(b) Load payment

Figure 8.13. Water Heater Energy Payments Using Optimal Control Strategies

8.5 Impact of Water Heater Capacity

The size of residential hot waters heater can range from 30 to 80 gallons. A larger water tank can store more heat thereby can coast longer when its set point is controlled (lowered). To access the affect of the electric water heater size on demand response a series of simulation where conducted for three demand response options: load curtailment, set point control, and modified set point control. The simulation results are shown in Figures 8.14, 8.15, and 8.16. There are several observations based on the results:

1. As expected, a water heater with a larger tank shows a lower morning peak demand. Assuming same amount of hot water consumption, a water heater with a larger tank turns on less frequently.

2. Although one would expect the coasting period to be longer because of additional storage capacity for a larger water heater, it was not the case. The reason is because the hot water usage is low during afternoon hours, when the demand response options are exercised.
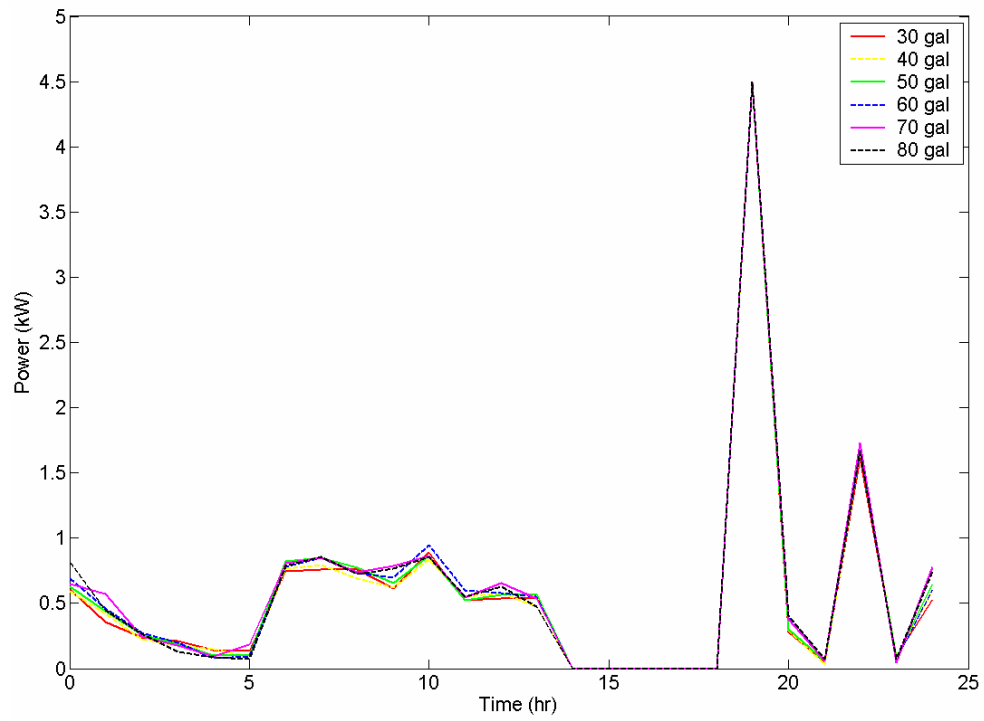
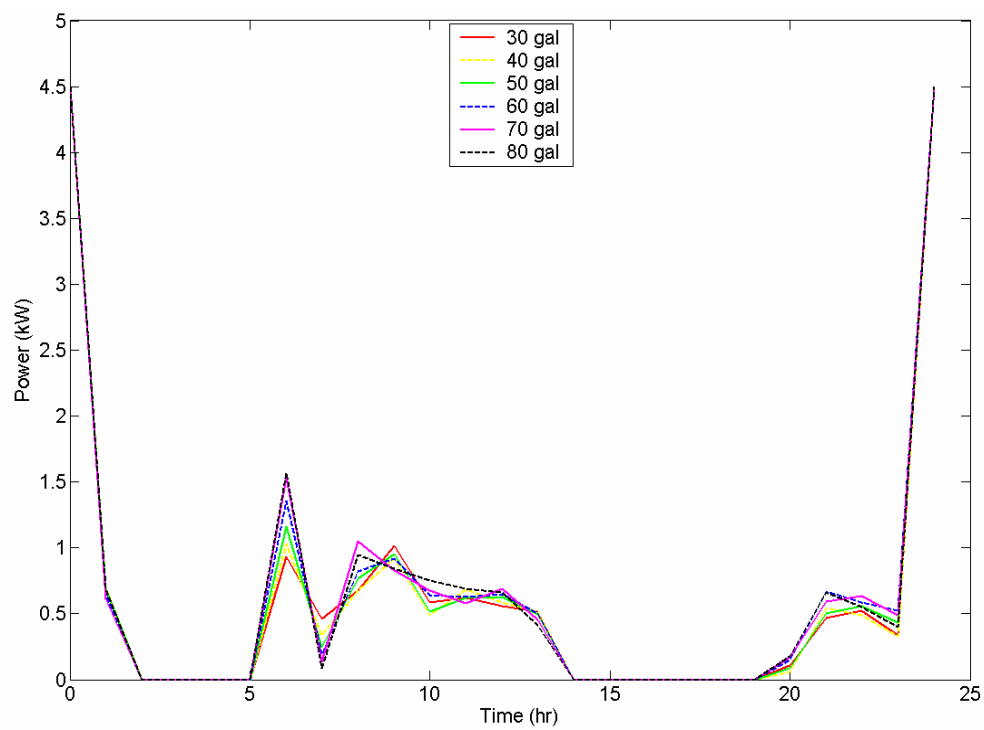**Figure 8.14.   Case 1: Water Heater Load Curtailment Option**

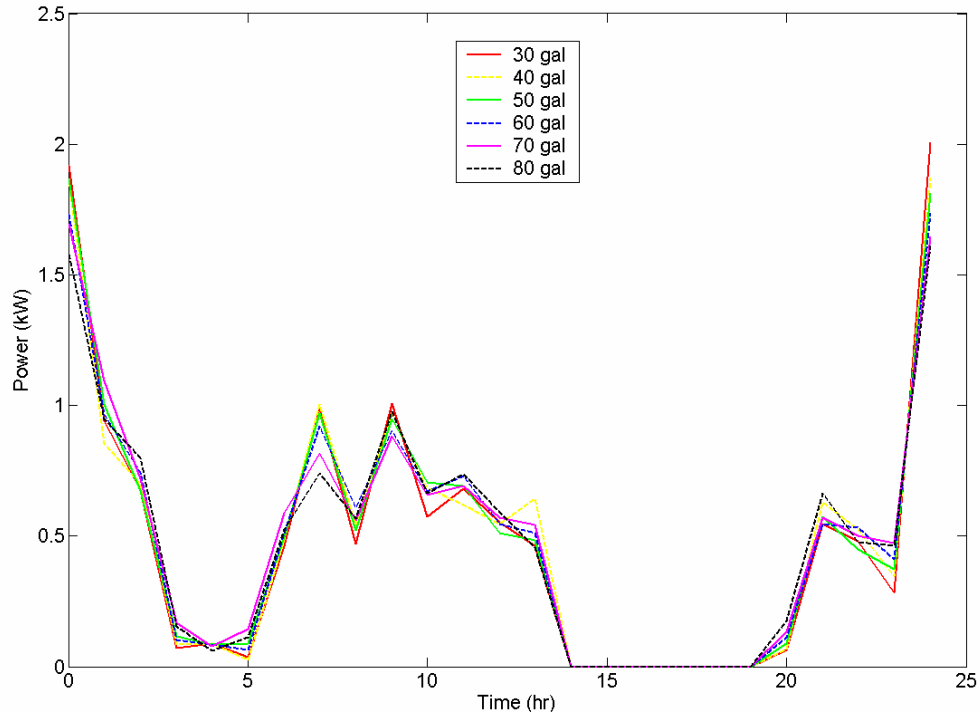**Figure 8.15.   Case 2: Water Heater Set Point Control Option**

Figure 8.16.   Case 3: Water Heater Modified Set Point Control Option

8.6   Evaluation of Control Strategies for HVAC Systems

This section focuses on developing and evaluating the control strategies for the HVAC systems using a simplified model based on first principles. As discussed earlier, there are several control strategies that can make TCA loads more elastic. In this section, four such control strategies will be evaluated: 1) load curtailment, 2) temperature reset, 3) modified temperature reset, and 4) precooling. Unlike hot water heaters where power consumption was the primary variable of interest, for HVAC systems in addition to the power consumption the comfort of the occupants also has to be monitored.

For the evaluation and comparison of the various control strategies, two base cases were simulated, the first one with an indoor temperature set point of 75°F and the second of 80°F. The same electricity price profiles used for simulating the hot water heater control strategies are used here (Figure 8.12). The peak price periods are generally between 1 p.m. and 6 p.m.   Power consumption of the two base cases, load curtailment controls, and the two reset control strategies is shown in Figure 8.17. The following observations are based on the simulation results:

1. Load curtailment option (blue line) results in most energy reduction during the peak price period resulting in more savings than the two reset control strategies. However, the room air dry-bulb temperature during the curtailment period is above the occupant comfort level (Figure 8.18). The room air dry-bulb temperature, which represents customer thermal comfort for the most part during the curtailment, is above 85°F. Because this control strategy leads to significant increase in room air dry-bulb temperature, people are less likely to implement a "hard" load curtailment of prolonged periods (greater than 1 hour).
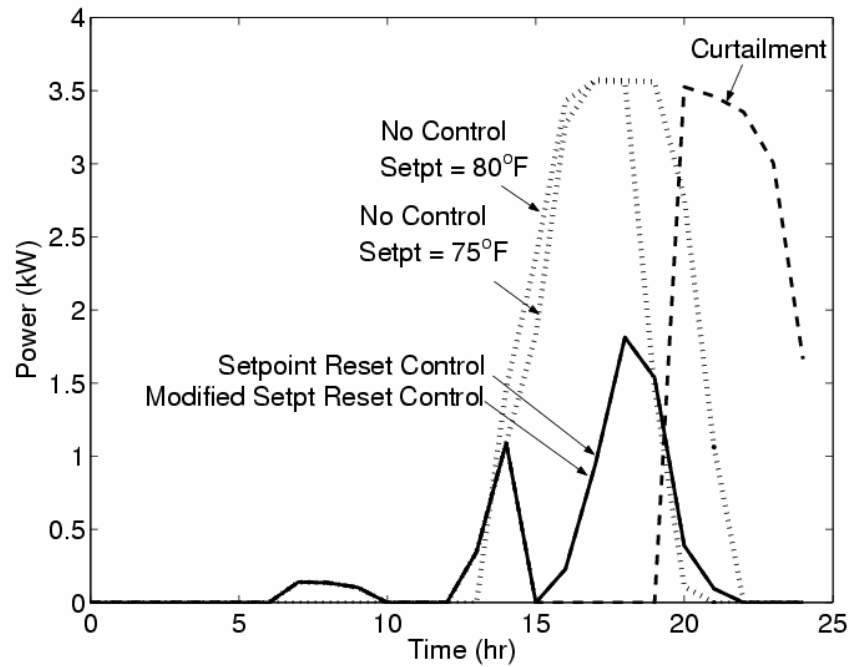
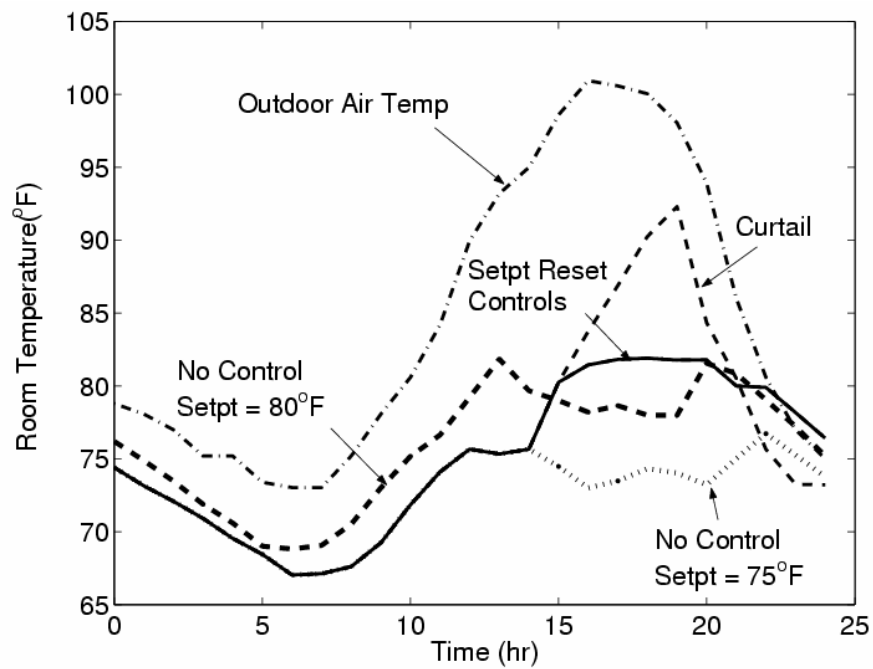**Figure 8.17. Power Consumption of the HVAC System under Temperature Reset Control**

Figure 8.18. Variation of Room Air Temperature under Different Control Strategies and the Outdoor Air Dry-Bulb Temperature
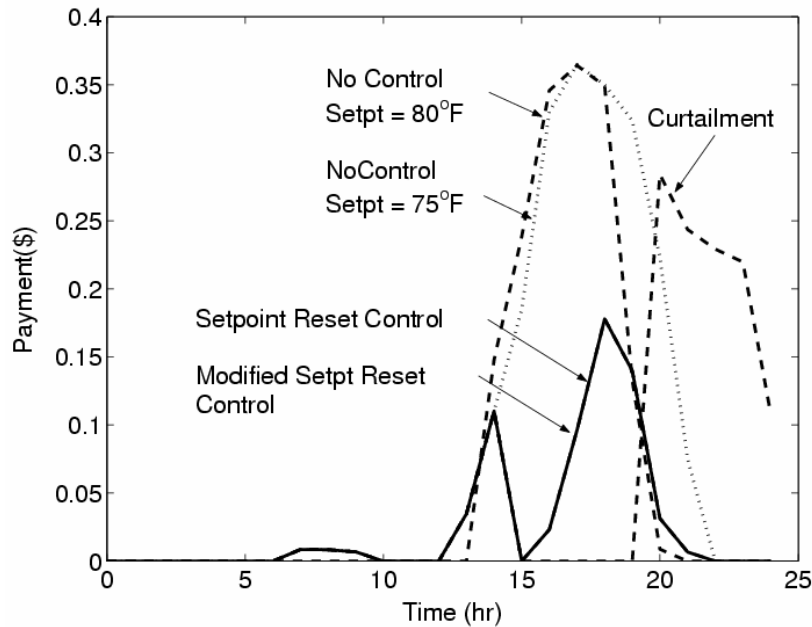
Figure 8.19.   Energy Payment under Different HVAC System Control Strategies

2. There are two base cases as noted earlier. As expected, if the base case set point is higher, the savings from load control options are smaller.

3. The second control strategy studied was the set point reset strategy. In the set point reset strategy, the set point of the residence is increased above the baseline set point during peak price periods (1 p.m. to 6 p.m.). The set point remains 75°F through 12 p.m. and is raised to 80°F at 1 p.m. when the peak price period begins. The reset control strategy results in fewer saving than the "hard" load curtailment strategy because the HVAC system is not completely turned off, unlike in the case of load curtailment. The HVAC system during temperature reset turns on less frequently than base case because the set point is raised. Unlike "hard" load curtailment, the indoor air-temperature is always below 80°F, which is within customer thermal comfort limits. Therefore, occupants may prefer this control strategy even for prolonged periods (greater than 1 hour).

4. The third control strategy studied was the modified set point reset strategy. In the modified set point reset strategy, the set point of the residence is increased above the baseline set point during peak price periods just like in the normal set point reset strategy. However, after the peak price period is over, the set point is gradually decreased to 75°F in 1°F/hr steps. The modified temperature reset strategy will result in similar energy savings and customer thermal comfort as the normal temperature reset during the curtailment period. We expected to see the modified set point control consume less power immediately after 6 p.m., which would have restored diversity if many home owners chose to employ the same control strategy. However, there was no change between set point reset and modified set point reset. It is possible that rate of change (1°F/hr) was large compared to the thermal mass in the house. Additional simulation with smaller rate of change should be studied.

5. The daily energy payment under different control strategies are shown in Figure 8.19 and tabulated in Table 8.3.

Table 8.3.   Energy Payment under Different Control Strategies

Control Strategies	Payment ($/day)
Setpoint 75°F	2.02
Setpoint 80°F	1.59
Curtailment	1.26
Pre-cooling	0.64
Modified Pre-cooling	0.64

Table 8.4.   Energy payments under Different Control Strategies for Combined HVAC and Hot Water Heater Simulation

Control Strategies	Payment ($/day)
Setpoint 75°F	6.69
Curtailment	5.50
Pre-cooling	5.19
Modified Pre-cooling	4.69

8.7   Control Strategies for Multiple TCAs

The power consumption, room air temperature, and the energy payment when applying curtailment, temperature reset, modified temperature reset control strategies of a mixed load that consists of HVAC system and electric water heater unit are shown in Figure 8.20 and tabulated in Table 8.4.

Based on the results, we have the following observations:

1. Curtailment (from 1 p.m. to 6 p.m.) to (blue lines in the figures) will result in the most energy reductions during the curtailment period and the most savings over the other type of control strategies. The disadvantages are the high payback load peaks and the sacrifice of customer comfort.

2. The pre-cooling/heating set point-control-strategy (pink lines in the figures) is implemented by pre-cooling HVAC units and pre-heating water heater units at 11 p.m. when the price is low, and raise/decrease the thermostat set points to cruise when the price is peaking.
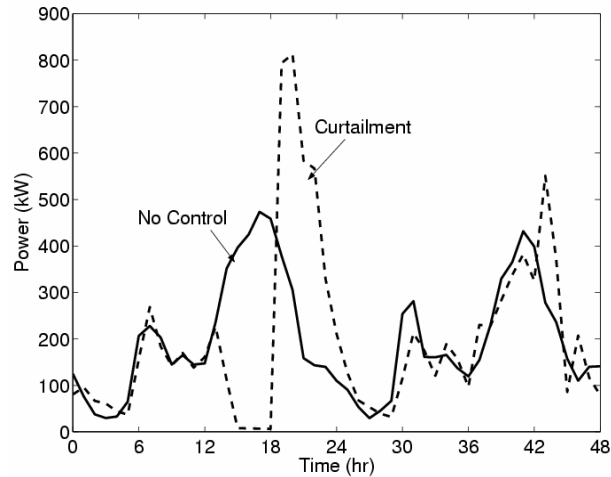
(a) Curtailment vs No-control
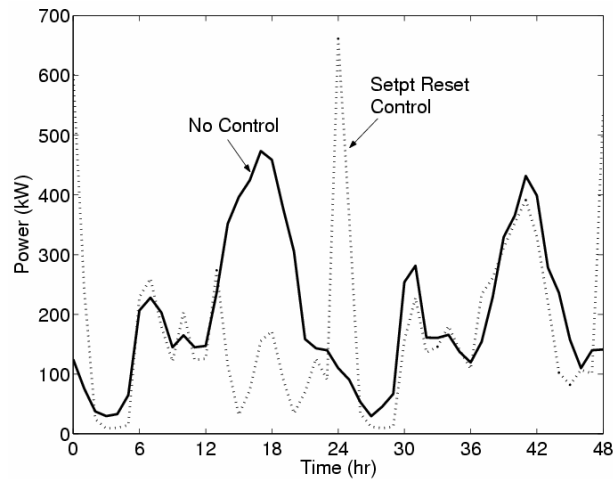
(b) Set Point Reset Versus No-Control
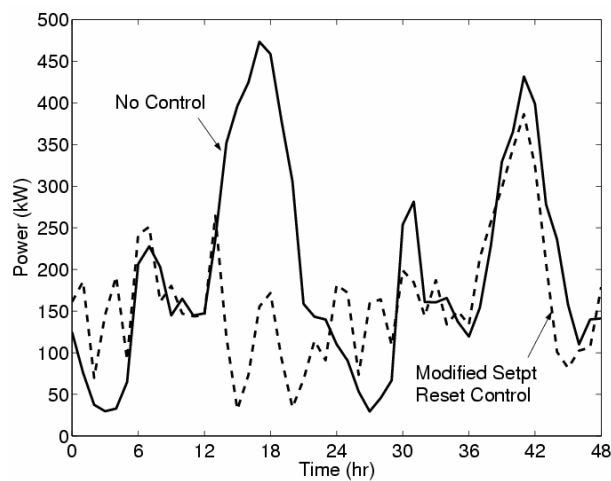
(c) Modified Set Point Reset Versus No-Control

Figure 8.20. Power Consumption under Different Control Strategies for Combined HVAC and Water Heater

9.0   Conclusions

We have considered the consequences of using concepts adopted from statistical mechanics in Part 1 as they relate to the response of highly distributed energy systems to bulk power markets. We defined how a market interacts with the system and rigorously derived its minimal cost configuration. We have observed that the transition to the super distributed energy system has inherent obstacles, which we have quantified.

We have seen that the Ising model of highly distributed energy systems reveals the existence of an important phase transition in the behavior of these systems when topological considerations are included. Below a critical level of trading activity in the system, the system cannot indefinitely sustain any distributed energy behavior at all, and only at or above the critical level of activity can the system engage in an effective market-based control of distributed resources.

The assumptions made in the derivation have placed restrictions on the conditions for which these results are accurate. For example, we have assumed configurations of the system that preclude the use these models for extreme conditions such as those observed with trading activity is very low with respect to the value of the trades. Despite the utility of the Ising model, we recommend investigation of the Potts model to account for other possible states that customer equipment may occupy, such as failed or withheld. Finally, we recommend further analysis of the drop in correlation distance in very active highly distributed energy systems because the Ising model suggests that under certain conditions, the behavior of market-based control systems may become largely insensitive to the latency of the communication network.

Two different modeling methods (SQ and ETP) were developed and used to evaluate various demand response control strategies for both electric water heaters and HVAC systems. In addition, we also studied some implementation issues. Although the results are promising in many aspects, more widely spread analysis is required to gain further insight into the affects of load curtailment on the distribution system dynamics.

In Section 7 developed a standby state queueing model to simulate the price response of a load consisting of thermostatically controlled appliances in a competitive electricity market. An aggregated load consists of thousands of TCAs, while the number of states a TCA can reside in may be no more than 100. We expect that applying a queue representation brings a computational advantage over simulating the behavior of each individual unit.

By analyzing the load shifts caused by the set point changes in response to price, the impacts on load diversity was studied. The results reveals the fundamental reason for the reduced load diversity in large scale direct load control systems, which has been observed and discussed by Weller (1988). The results also indicate that by responding to price changes, a diversified TCA type of load becomes synchronized, and their behaviors present a dynamic response. Therefore, to design a successful load response program for aggregated TCAs, one needs to examine the load shifting characters to ensure that the shifted load peaks will occur after the peak-price time. The synchronized load peak can be much higher than that of the diversified load. The stress on the distribution system should also be considered. The methodology developed in this research is expected to be used to create DSM simulation tools that are able to take the load-shifting behavior into consideration.

In Section 6 we presented a method to include the modeling of uncertainties in the TCA load cycling times and random load consumptions by modifying the entries of the state transition matrix of a SQ model. From the load survey data, probability distribution functions of the cycling time and consumer behavior patterns can be obtained. State transition matrixes for each hour can then be tuned and used by the SQ model introduced in (Lu et al. 2004c) to simulate the price response of loads aggregated by TCAs. The eigenvalues of the state transition matrix can be used to evaluate the system damping rate, which indicates the ability of the system to restore load diversity. Because the computation time of a SQ model is determined by the number of states used to model the TCA load cycling periods and the sparsity of the state transition matrix, it has computational advantages over unit-based simulation, thus holding promise as an useful equivalent load model for transmission level studies.

In Section 7 we investigated several control strategies including an optimal control strategy for an LSE to implement TCA set point-control in a competitive electricity market, where the market clearing price is insensitive to the bid price of a single load bid. A state queueing model has been used to simulate the aggregated water heater load response. Economic benefits are calculated and feeder load profiles are obtained to evaluate control strategies including load curtailment, preheating and coasting, and modified preheating and coasting. The results suggest that the modified preheating and coasting control is cost saving and result in least dynamics on feeder load profile.

Future work will be focused on developing control strategies to minimize the energy cost in real time market. The possibility of using TCAs to provide ancillary service, where the load needs to respond to frequency deviations, is also going to be investigated.

10.0   References

Abowd, J. 2003. personal communication J. Malard.

Albert, R. et al. 2000. "Error and attack tolerance of complex networks," Nature, 406:378-381.

Alvarado, FL and R Rajaraman. 2000. "Understanding price volatility in electricity markets." Proceedings of the 33rd Annual Hawaii International Conference on System Sciences, Jan.

Alvarado, F. 2003. "Is system control entirely by price feasible?" in proc. of the 36th Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos, California.

Armstrong, PR et al. 2001. "Controls," in Handbook of Heating Ventilation and Air Conditioning, J. F. Kreider ed., CRC Press, New York.

American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc. 2001. ASHRAE HandbookCD.

Atay, FM. 2004. "Total and partial amplitude death in networks of diffusively couple oscillators," arXiv, http://arxiv.org/pdf/nlin/0401033.

Atmanspacher, H and G Wiedenmann. 1999. "Some basic problems with complex systems," in Large Scale Systems: Theory and Applications, ed. by N. T. Koussoulas and P. Groumpos, Elsevier Publ., Amsterdam, pp. 1059-1066.

Bak, P, C Tang, K Wiesenfelf. 1988. "Self-organizing criticality," Physical Review A:38,364.

Bak, P and K Sneppen. 1993. "Punctuated equilibrium and criticality in a simple model of evolution," Physical Review Letters 71, 4083-4086.

Bar-Yam, Y. 2000. Concepts in Complex Systems. Available URL: http://necsi.org/guide/concepts/statisticalmechanics.html.

Barabasi, A-L. 2002. Linked: The New Science of Networks, Perseus Press, Cambridge, Massachussetts.

Bar-Yam, Y. 1997. "Dynamics of Complex Systems (Studies in Nonlinearity)", Available URL: http://necsi.org/publications/dcs/index.html.

Bauer, S, TL Henderson, S Summers, and G Brown. 2001. "Refrigerator and Freezer Safety." Available URL: http://www.wvu.edu/~exten/infores/pubs/fypubs/136.wl.pdf.

Bertsekas, DP. 1992. "Auction Algorithms for Network Flow Problems: A Tutorial Introduction," Computational Optimization and Applications, vol. 1, pp. 7-66.

Bohn, RE. 1982. "Spot Pricing of Public Utility Services", Ph.D. dissertation, MIT, 1982. Also available as MIT Energy Laboratory Technical Report, MIT-EL 82-031.

Buchanan, M. 2002. Nexus: Small Worlds and the Groundbreaking Science of Networks, W. W. Norton & Co., New York.

Carvalho-Rodrigues, F, J Dockery, and AER Woodcock. 1992. "Models of combat with embedded command and control: casualty based entropy calculations as a combat predictor," in The Military Landscape, Mathematical Models of Combat, Woodhead Press, Cambridge, UK.

Carvalho-Rodrigues, F, J Dockery, and T Rodrigues. 1993. "Entropy of plagues: A measure assessing the loss of social cohesion due to epidemics," European Journal of Operational Research 71, 45-60.

Chassin, DP. 2003. "Power market dynamics: The statistical mechanics of transaction-based control," in proc. of ECOS 2003, Copenhagen, Denmark.

Chassin, DP. 2004. "Statistical Mechanics: A Possible Model for Market-based Systems Dynamics", in proc. 37th Annual Hawaii International Conference on System Sciences.

Chassin, DP. 2004. "Properties and Phases of Highly Distributed Energy Systems Under Decentralized Autonomous Control (Parts 1 & 2)", *IEEE Transactions in Power Systems*, submitted August [in review].

Chan, ML, EN Marsh, JY Yoon, GB Ackerman, and N Stoughton. 1981. "Simulation-based Load Synthesis Methodology for Evaluating Load-management Programs." IEEE Trans. on Power Apparatus and Systems, vol. PAS-100, pp. 1771 - 1778.

Clausius, RJE. 1865. "Die mechanische Wärmetheorie", *Annalen der Physik*, *Series 2*, vol. 125, no. 426 [in German].

Cont, R. 1999. "Modeling economic randomness: statistical mechanics of market phenomena," in Statistical Physics in the eve of the 21st Century: the James B. McGuire Festschrift, M. T. Bachelor & L. T. Wille eds., World Scientific, Singapore.

Cooper, F, A Khare, U Sukhatme. 1995. "Supersymmetry and Quantum Mechanics", Phys. Rept. 251, pp. 267-385.

Daryanian, B, RE Bohn, and RD Tabors. 1989. "Optimal Demand-side Response to Electricity Spot Prices for Storage-type Customers," IEEE Trans. on Power Systems, vol. 4, pp. 897–903.

Dexter, P. 2003. "Combat entropy as a measure of effectiveness," Journal of Battlefield Technology, Vol. 6, No. 3.

Donangelo, R and K Sneppen. 2003. "Self-organization of value and demand," Physica A 276:572-580.

Dorogovtsev., SN, AV Golstev, and JFF Mendes. 2003. "Potts model on complex networks", arXiv cond-mat 0310693.

Dutt, R, A Gangopadhyaya, A Khare, and UP Sukhatme. 1994. "Thermodynamics of a Free Q-Fermion Gas", International Journal of Modern Physics A, vol. 9, no. 15, pp. 2687-2698.

Edwards, FV, GJW Dudgeon, JR McDonald, WE Leithead. 2000. "Dynamics of Distribution Networks with Distributed Generation", IEEE Power Engineering Society Summer Meeting, vol. 3, pp. 1032-1037.

Eto, J, C Marnay, C Goldman, J Kueck, J Dagle, F Alvarado, T Mount, S Oren, and C Martinez. 2001. "An R&D Agenda to Enhance Electricity System Reliability by Increasing Customer Participation in Emerging Competitive Markets", Power Engineering Society Winter Meeting, vol. 1, pp.247-251.

Feng, J, et al. 2000. "Convergence to global minima for a class of diffusion processes," Physica A 276:465-476.

Feynman, R. 1972. *Lectures on Statistical Mechanics*. Addison-Wesley.

Fleay, BJ. 2001. "Background paper November 2001", in report of Electricity Reform Task Force, Government of Western Australia, unpublished.

Focardi, SM and FJ Fabozzi. 2004. The Mathematics of Financial Modeling and Investment Management, Wiley Finance, John Wiley & Sons, New Jersey, p.78.

Foley, DK. 2002. "Maximum Entropy Exchange Equilibrium", Department of Economics, Barnard College, Columbia University URL http://cepa.newschool.edu/~foleyd/maxentexeq.pdf.

Freeman, LC. 2000. "Visualizing Social Networks," Journal of Social Structure, vol. 1, no. 1, Carnegie Mellon University, Pittsburg, Pennsylvania.

Fuhrman, S, MJ Cunningham, X Wen, G Zweiger, JJ Seihamer, R Somogyi. 2000. "The application of Shannon entropy in the identification of putative drug targets," Biosystems 55, 5-14.

Gellings, CW and RW Taylor. 1981. "Electric Load Curve Synthesis – A Computer Simulation of an Electric Utility Load Shape" IEEE Trans. on Power Apparatus and Systems, vol. PAS-100, No. 1, Jan.

Gell-Mann, M. 1994. The Quark and the Jaguar, W. H. Freeman & Co., New York.

Georgescu-Roegen, N. 1971. The Entropy Law and the Economic Process, Harvard Univ. Press, Cambridge, Massachussetts.

Ghosh, AK, DL Lubkeman, MJ Downey, and RH Jones. 1997. "Distribution Circuit State Estimation Using a Probabilistic Approach." *IEEE Transactions on Power Systems*, vol. 12, pp. 45–51.

Gibbs, JW. 1902. *Elementary Principles in Statistical Mechanics*. New York, Dover Publications, 1960.

Glosh, AK, DL Lubkeman, MJ Downey, and RH Jones. 1997. "Distribution Circuit State Estimation Using a Probabilistic Approach", *IEEE Transactions on Power Systems*, vol. 12, no. 1, pp. 45-51.

Grigsby. 1999. Power Systems Engineering Handbook, IEEE Press, Wiley & Sons, p.9-49

Gross, D and CM Harris. 1998. Fundamentals of Queueing Theory, 3rd ed., John Wiley & Sons. Inc.

Guttromson, RT. 2002. "Modeling Distributed Energy Resource Dynamics on the Transmission Systems", IEEE Transactions on Power Systems, vol. 17, no. 4, pp. 1148-1153.

Guttromson, RT, DP Chassin, and SE Widergren. 2003. "Residential Energy Resource Models for Distribution Feeder Simulation." Proc. of 2003 IEEE PES General Meeting, Toronto, Canada, pp. 108-113.

Hatziargyriou, ND, TS Karakatsanis, and M Papadopoulos. 1990 "Probabilistic Calculations of Aggregate Storage Heating Loads." IEEE Trans. on Power Delivery, vol. 5, No. 3, pp. 1520 – 1526.

Hauer, JF et al. 2001. "Direct analysis of wide area dynamics," in The Electric Power Engineering Handbook, LL Grigsby (ed.), IEEE Press, Boca Raton, Florida.

Henderson, JM and RE Quandt. 1980. *Microeconomic Theory: A Mathematical Approach*, 3rd Edition, McGraw Hill, New York.

Herman, M. 1999. "Modeling the revolution in military affairs," Joint Force Quarterly, Autumn/Winter 1998-99, 85-90.

HowStuffWorks Inc. 2004. Available URL:http://www.howstuffworks.com/water-heater.htm.

Huberman, B and S Clearwater. 1995. "A multi-agent system for controlling building environments," in proc. of the International Conference on Multiagent Systems, pp. 171-176.

Hubbard, BB. *The World According to Wavelets*, 2nd edition, A. K. Peters, Ltd. Natick, Massachusetts, pp. 281-283.

HyungSeon, O, RJ Thomas, B Leiseutre, TD Mount. 2001. "Comparing the behavior of agents to human subjects in a uniform price auction," in proc. of 36th Annual Hawaii International Conference on System Sciences, IEEE Computer Society, 2003.A.-M. Borbely and J. F. Kreider, "Distributed Generation: The Power Paradigm for the New Millenium", CRC Press.

Ihara, S and FC Schweppe. 1981. "Physically Based Modeling of Cold Load Pickup." IEEE Trans. on Power Apparatus and Systems, vol. PAS-100, No. 9, Sept.

Ising, E. 1925. "Beitrag zur Theorie des Ferromagnetismus", *Zeitschr. f. Physik*, vol. 31, pp 253-258, (german).

Istrail, S. 2000. "Statistical mechanics, three-dimensionality and NP-completeness: I. Universality of intracatability for the partition function of the Ising model across non-planar surfaces", in proc. of *Thirty-Second Annual ACM Symposium on Theory of Computing* (STOC 2000), pp. 87-96.

James, J and F Mabry. 2004. "Building trustworthy systems: guided state estimation as a feasible approach for interpretation, decision and action based on sensor data," in proc. of the 37th Hawaii International Conference on System Sciences, IEEE Computer Society, Los Alamitos, California.

Jaynes, ET. 1991. *How should we use entropy in economics?*, unpublished manuscript originally written in 1982, last revised in 1991, URL http://bayes.wustl.edu/etj/articles/entropy.in.economics.pdf.

Jaynes, ET. 1986. "Predictive statistical mechanics," in Frontiers of Nonequilibrium Stastitical Physics, G. T. Moore and M. O. Scully (eds.), Plenum Press, New York, p.33.

Jeffrey, A. 2004. *Mathematical and Analytical Methods in the Physical Sciences: Inverse Problems and Singular Perturbation Theory*, Volume 1, Kluwer Academic Publishers, Dordrecht, Holland.

Johnson, NL and FC Leone. 1977. Statistics and Experimental Design in Engineering and Physical Science. John Wiley & Sons, Ltd., New York, 1977.

Kannberg, LD, DP Chassin, JG Desteese, SG Hauser, MC Kintner-Meyer, RG Pratt, LA Schienbein, WM Warwick. 2003. *GridWiseTM: The Benefits of a Transformed Energy System*, Report no. PNNL- 14396, Pacific Northwest National Laboratory, Richland, Washington, September.

Kaneko, K and T. Yomo. 1997. Bull. Math. Biol. 59, 139.

Kaufman, S. 2000. Investigations, Oxford University Press, Oxford.

Kaufman, B and L Onsager. 1953. "Crystal statistics. III. Short-range order in a binary Ising lattice", *Physical Review*, vol. 76, pp. 1244-1252.

Kittel, C. 1969. *Thermal Physics*, 1st edition, John Wiley & Sons, New York.

Kramers, HA and GH Wannier. 1941. "Statistics of Two-Dimenstional Ferromagnet, Part I", *Physical Review*, vol. 60, 1941, pp. 252-262; H. A. Kramers and G. H. Wannier, "Statistics of Two-Dimenstional Ferromagnet, Part II", *Physical Review*, vol. 60, pp.263-276.

Kummer, A and R Ocone. 2003. "Extension of dynamics of granular flow methodology to cell biology," Physica A 321, 587-597.

Kurzyński, M. 2000. "Internal dynamics of biomolecules and statistical theory of biochemical processes," Physica A 285, 29-47.

Langmean, CJ, CR McClung, BR Donald. 2002. " A maximum entropy algorithm for rhythmic analysis of genome-wide expression patterns," in proc. of CSB'02, *IEEE*, August.

Lee, SH and CL Wilkins. 1983. " A Practical Approach to Appliance Load Control Analysis: A water Heater Case Study." IEEE Trans. on Power Apparatus and Systems, vol. PAS-102, pp. 1007–1012, April.

Leichti, H. 2000. Summary of Laws: Thermodynamics, Kepler, Newton, Enstein, http://www.geocities.com/ciencia_farma/nat_laws.htm.

Lenz, W. 1920. "Beitrag zum Verstandnis der magnetischen Eigenschaften in festen Korpern", *Phys. Zeitschr*, vol. 21, pp. 613-615.

Lu, N and DP Chassin. 2004. *IEEE Power Systems*, 19 3.

Lu, N, DP Chassin, and SE Widergren. 2004. "Simulating Price Responsive Distributed Energy Resouces." In Proc. of the IEEE Power System Conference and Exposition 2004, NYC, New York. [In Press].

Lu, N, DP Chassin, and SE Widergren. 2004. Modeling Uncertainties in Aggregated Thermostatically Controlled Loads Using a State Queueing Model. PNNL-SA-41229, Pacific Northwest National Laboratory, Richland, Washington.

Lu, N and S Katipamula. 2004. Control Strategies of Thermostatically Controlled Appliances in a Competitive Electricity Market. PNNL-SA-*****, Pacific Northwest National Laboratory, Richland, Washington.

Lu, N, JH Chow, and AA Desrochers. 2004. "Pumped-Storage Hydro-Turbine Bidding Strategies in a Competitive Electricity Market." IEEE Trans. on Power Systems, vol. 19, pp. 834-842, May.

Molina, A, A Gabaldon, JA Fuentes, and C Alvarez. 2003. "Implementation and Assessment of Physically Based Electrical Load Models: Application to Direct Load Control Residential Programmes." Generation, Transmission and Distribution, IEE Proceedings-, vol. 150, pp. 61 - 66.

Moreno, Y and AF Pacheco. 2004. "Synchronization of phase oscillators in scale-free networks," arXiv, http://arxiv.org/pdf/cond-mat/0401266.

Morowitz, H. 1992. Beginnings of Cellular Life: Metabolism Recapitulates Biogenesis. New Haven: Yale University Press.

Nazarko, J, RP Broadwater, and NI Tawalbeh. 1998. "Identification of Statistical Properties of Diversity and Conversion Factors from Load Research Data." Electrotechnical Conference, MELECON 98., 9th Mediterranean, vol. 1, pp. 217 - 220.

Newman, MEJ. 2000. "Small worlds: The structure of social networks," Complexity.

Nemenman et al. 2003. "Entropy and information in neural spike trains: Process on the sampling problem," arXiv, http://www.arXiv.org/physics/0306063.

NYISO. www.nyiso.com.

Oliveira, JS, JB Jones-Oliveira, DP Chassin. 2003. "A computational model for the identification of critical pathways in an electric power grid with distributed generation," PNNL-SA-38403, Pacific Northwest National Laboratory, Richland, Washington. [unpublished]

Oliveira, JS, CG Bailey, JB Jones-Oliveira, DA Dixon, DW Gull, and ML Chandler. 2003. "A computational model for the identification of biochemical pathways in the Krebs cycle," J. Comput. Biol. 10(1), pp. 57-82.

Onsager, L. 1944. "Crystal statistics. I. A two-dimensional model with an order-disorder transition", *Physical Review*, vol. 65, pp. 117-149.

Pahwa, A and CW Brice III. 1985. "Modeling and System Identification of Residential Air Conditioning Load." IEEE Trans. on Power Apparatus and Systems, vol. PAS-104, No. 6, June.

Patel, A. 2000. "Quantum Algorithms and the Genetic Code," Lecture in proc. of Winter Institute on Foundations of Quantum Theory and Quantum Optics, 1-13 January 2000, *S. N. Bose National Centre for Basic Sciences*, Calcutta, India, http://www.arXiv.org/quant-ph/0002037.

Peierls, R. 1936. "On Ising's Model of Ferromagnetism", *Proceedings of the Cambridge Philosophical Society*, vol. 32, pp. 477-481 (in german).

Perc, M and M. Marhl. 2004. "Frequency dependent stochastic resonance in a model for intracellular Ca^{2+} oscillations can be explained by local divergence," Physica A 332, 123-40.

Peterson, FJ, JE Patton, ME Miller, RA Gillman, WM Warwick, and WF Sandusky. 1993. "End-Use Load and Consumer Assessment Program." Energy and Buildings, ISSN 0378-7788, vol. 19, No. 3.

Pindyck, RS and DL Rubinfeld. 2000. Microeconomics (5th Edition). Pearson Education.

Pratt, RG, CC Conner, EE Richman, KG Ritland, WF Sandusky, ME Taylor. 1989. "Description of Electric Energy Use in Single-family Residences in the Pacific Northwest," *End-use Load and Consumer Assessment Program (ELCAP)*, report to Bonneville Power Administration, Pacific Northwest National Laboratory, Richland, Washington, July.

Pratt, RG and ZT Taylor. 1994. "Development and Testing of an Equivalent Thermal Parameter Model of Commercial Buildings from Time-Series End-Use Data." Pacific Northwest Laboratory, Richland, WA, Apr.

Ptashne, M. 1992. A Genetic Switch, (Second ed.), Cell Press, Cambridge, Massachussetts.

Roddestvenski, I and MG Cottam. 2000. "Time rescaling and generalized entropy in relation to the internal measurement concept," Physica A 278, 150-160.

Ryan, NE, SD Braithwait, JT Powers, and BA Smith. 1989. "Generalizing Direct Load Control Program Analysis: Implementation of the Duty Cycle Approach." IEEE *Transactions on Power Systems*, vol. 4, pp. 293 – 299.

Schaden, M. 2004. "Quantum Finance", *Physica A* 316:1-4, pp. 511-538.

Schweppe, FC, B Daryanian, and RD Tabors. 1989. "Algorithms for a Spot Price Responding Residential Load Controller." IEEE Transactions on Power Systems, vol. 4, pp. 507–516, May.

Sergeev, V. 2001. "The thermodynamic approach to market equilibrium," Santa Fe Institute Working Papers, Santa Fe, New Mexico, Nov.

Sergeev, V. 2003. "The Thermodynamic Approach to the Market Equilibrium", Santa Fe Institute Working Papers, 03-04-027.

Sergeev, V. 2003. "The Thermodynamic Approach to the Market Equilibrium", *Santa Fe Institute Working Papers*, 03-04-027.

Sergeev, V. 2003. The Thermodynamic Approach to Market Equilibrium, Santa Fe Working Papers, April.

Shafee, F. 2004. *Chaos and Annealing in Social Networks*, arXiv, http://arXiv.org/pdf/cond-mat/0401191.

Silva, HS and ML Martins. 2003. "A cellular automata model for cell differentiation," Physica A 322, 555-566.

Smith, A. 1776. *The Wealth of Nations*.

Smith, E and D Foley. 2002. Is utility theory so different from thermodynamics, Santa Fe Working Papers, SFI#02-04-016, Santa Fe Institute, Santa Fe, New Mexico.

Smith, R. 1980. "The Contract Net Protocol: high-level communication and control in a distributed problem solver," IEEE Transactions on Computers, C-29(12):1104–1113.

Stoft, S. 2002. *Power System Economics: Designing Markets for Electricity*, IEEE Press, Boca Raton, 2002.

Taylor, ZT and RG Pratt. 1988. "The Effects of Model Simplifications on Equivalent Thermal Parameters Calculated from Hourly Building Performance Data." Proceedings of the 1988 ACEEE Summer Study on Energy Efficiency in Buildings, pp. 10.268 – 10.285.

Taylor, ZT, RG Pratt. 1989. "Description of Electric Energy End-use in Commercial Buildings in the Pacific Northwest", report no. DOE/BP-13795-22: *End-use Load and Consumer Assessment Program (ELCAP)*, Bonneville Power Administration.

Tsallis, C et al. 2003. "Nonextensive statistical mechanics and economics," Physica A 324:89.

Vilar, JMG et al. 2002. "Mechanisms of noise-resistance in genetic oscillators," in proc. of Natl. Acad. Sci. USA 99, 5988-5992.

Watts, DJ and SH Strogatz. 1998. "Collective dynamics of small-worlds networks," Nature, 393:440-442.

Weller, GH III. 1988. "Managing the Instantaneous Load Shape Impacts Caused by the Operation of a Large-scale Direct Load Control System." IEEE Trans. on Power Systems, vol. 3, pp. 197 – 199.

Westinghouse Electric Corporation. 1965. Electric Utility Engineering Reference Book – Distribution Systems, East Pittsburgh, Westinghouse Electric Corporation, vol. 3.

Wu, FY. 1982. *The Potts model*, Reviews of Modern Physics, vol. 54, pp. 235-268.

Yamamoto, Y, I Hidaka, D Nozaki, N Iso-o, R Soma and S Kwak. 2002. "Noise-induced sensitization of human brain," Physica A 314, 53-60.

Yasudo, K and T Ishii. 2002. "Decentralized Autonomous Control of Super Distributed Energy Systems", IEEE SMC.

Appendix A

Distributed Generation Systems

Appendix A

Distributed Generation Systems

The main shortcoming with the binary model of markets is that it does not accurately reflect the true nature of markets. The binary model assumes 1) that any participant can both sell and buy (although obviously not simultaneously), and 2) that the decision to buy or sell is 50/50 given an indifferent price. The reality is that end-use load cannot be generators, and only rarely have 50% duty cycles. Therefore it is necessary to employ a model of supply and demand that does not have this shortcoming. We need a model that allows us to determine the aggregate properties of a distribution system given knowledge of the number of distributed generators, the number of loads, and the probability that each will be on. In general we expect the load to be given by the equation

$$\langle Q \rangle = \left(N_{gen} \rho_{gen} - N_{load} \rho_{load} \right) q \tag{A.1}$$

where N_{gen} is the total number of generators, N_{load} is the total number of loads, ρ_{gen} is the probability that a generator is on, ρ_{load} is the probability that load is on, and q is the quantity of power transacted between a generator and a load.

We begin by considering only one side of the market. To derive the entropy of an arbitrary system of suppliers,[a] such as distributed generators willing to contract for the delivery of a quantity q of power, we propose to use Kittel's modified lattice gas model (Kittel 1969) of R available contracts, each of which is an agreement to transact one unit of q. We consider a very large number suppliers acting independently such that when they are contracted each accepts one transaction (a site in the lattice). Thus the uncontracted suppliers serve as a reservoir. Each transaction has zero suppliers when there is no contract or one supplier when a contract is accepted. We seek the probabilities

$$P(0), P(1), P(2), \cdots, P(n), \cdots \tag{A.2}$$

that a total of 0, 1, 2, … , n, … suppliers are engaged in the R available contracts, if we have an average number $\langle n \rangle$ of suppliers contracted over an ensemble of similar systems.

First we consider a system with a single contract. The grand sum for this system is

$$Z_1 = 1 + \lambda \tag{A.3}$$

where the term λ is proportional to the probability that the contract is accepted, and the term 1 is proportional to the probability that the contract is not. Thus the absolute probability that the contract is accepted is

$$f = \frac{\lambda}{1 + \lambda}. \tag{A.5}$$

(a) The same argument applies to consumers.

When $\lambda \ll 1$, then $f \cong \lambda$. The true value of λ is determined by the condition of the suppliers in the reservoir, because for diffusive contact between the lattice and the reservoir we always have

$$\lambda_{contracts} = \lambda_{suppliers} \tag{A.5}$$

according the arguments that led to the definition of the migration potential μ (Sergeev 2003). (The evaluation of the $\lambda_{suppliers}$ for an ideal system is given in Chapter 11 of Kittel 1969) We can extend this treatment to R contracts, for which

$$Z = Z_1 Z_2 \cdots Z_R = (1 + \lambda)^R . \tag{A.6}$$

The method used to employ the binomial expansion for the states of the spin model counts each and every state of the system of R contracts. Each contract has two alternative states, one active and the other vacant, which corresponds to the terms $1 = \lambda^0$ and $\lambda = \lambda^1$. In the low-limit occupancy of $f \ll 1$, we have $f \cong \lambda$, and so

$$\langle n \rangle = fR = \lambda R \tag{A.7}$$

is the average total number of active contracts. The probabilities we seek are concerned with this low-occupancy limit. We can now write equation (A.4) as

$$Z = \left(1 + \frac{\lambda R}{R}\right)^R = \left(1 + \frac{\langle n \rangle}{R}\right)^R . \tag{A.8}$$

If we let the number of contracts R increase without limit, while maintaining a constant average number of active contracts $\langle n \rangle$, then we have

$$\lim_{R \to \infty} \left(1 + \frac{\langle n \rangle}{R}\right)^R = e^{\langle n \rangle} \tag{A.9}$$

and expanding the exponential function as a power series

$$Z_R \cong e^{\langle n \rangle} = e^{\lambda R} = \sum_n \frac{(\lambda R)^n}{n!} . \tag{A.10}$$

The term λ^n is proportional to the probability $P(n)$ that n contracts are active. With the grand sum as the normalization factor we have in the limit of $R \to \infty$

$$P(n) = \frac{\lambda^n R^n}{n!} \frac{1}{Z_R} = \frac{\lambda^n R^n e^{-\lambda R}}{n!} . \tag{A.11}$$

With $\lambda R = \langle n \rangle$ we find that the probability $P(n)$ is the Poisson distribution

$$P(n) = \frac{\langle n \rangle^n e^{-n}}{n!} \tag{A.12}$$

and therefore, using Stirling's approximation the entropy is

$$\sigma(n) = \sigma_0 + n \log\langle n \rangle - \left(n + \tfrac{1}{2}\right)\log n - \tfrac{1}{2}\log 2\pi \tag{A.13}$$

as shown in Figure A.1.

The quantity σ_0 may be determined from the proportionality constant a required to make equation (A.12) a probability density function such that
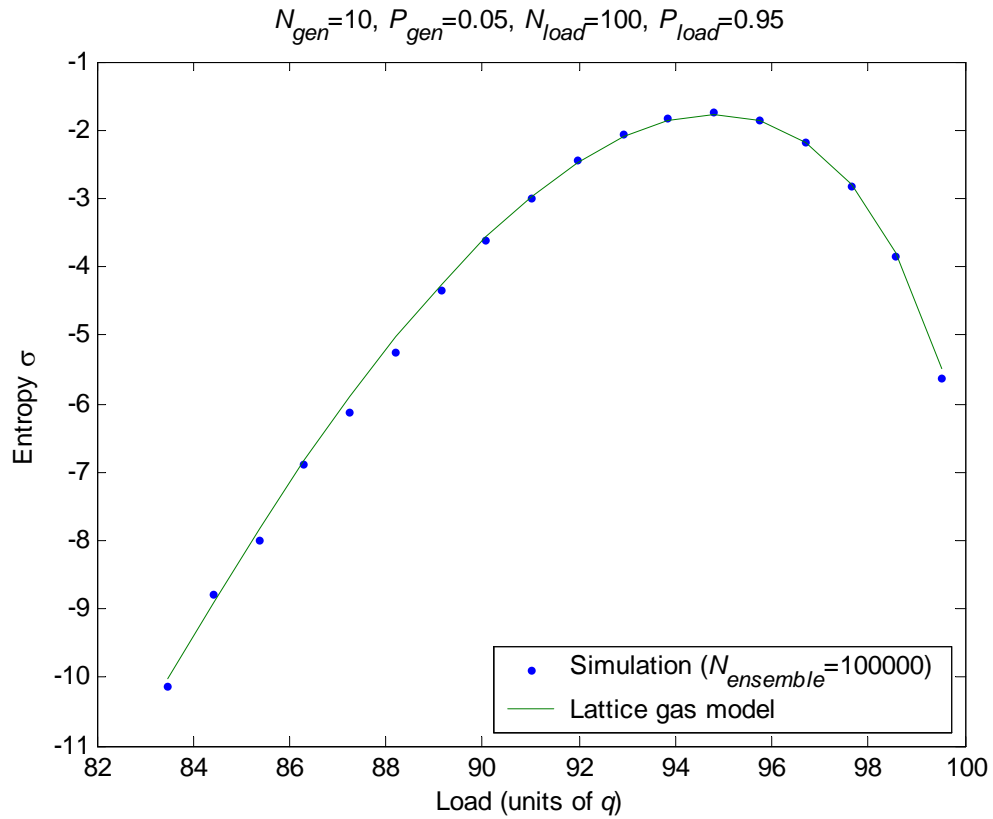

Figure A.1. Comparison of Simulated Distributed Generation and lattice Gas Model

$$a = \int_0^\infty P(x)\,dx \tag{A.14}$$

and thus $\sigma_0 = -\log a$. However, this value is generally not necessary because entropy differences are usually all that concern us.

To determine the equation of state for this system, we must establish how the value U of the system is determined. In this case we will treat the value as the revenue derived from satisfaction the unmet demand, i.e.,

$$U = QP = nqP = n\varepsilon .$$

(A.15)

The expectation value $\langle U \rangle$ of the system is similarly

$$\langle U \rangle = \langle Q \rangle P = \langle n \rangle qP = \langle n \rangle \varepsilon .$$

(A.16)

Therefore we find that

$$\sigma(U) = -\log a + \frac{U}{\varepsilon} \log \frac{\langle U \rangle}{\varepsilon} - \frac{1}{2} \log 2\pi - \left(\frac{U}{\varepsilon} + \frac{1}{2} \right) \log \frac{U}{\varepsilon}$$

(A.17)

From the definition of τ, we have

$$\frac{1}{\tau} = \frac{\partial \sigma}{\partial U} = \frac{U + \frac{1}{2}\varepsilon}{U} + \frac{1}{\varepsilon} \log \frac{\langle U \rangle}{\varepsilon}$$

(A.18)

or

$$U = -\frac{1}{2} \frac{\varepsilon \tau}{\left(1 + \frac{1}{\varepsilon} \log \langle n \rangle \right) \tau - 1} .$$

(A.19)

The denominator causes a discontinuity in U at the trading activity

$$\tau_c = \frac{\varepsilon}{\varepsilon + \log \langle n \rangle}$$

(A.20)

below which the value U is positive (i.e., the system imports power) and above which the value of negative (i.e., the system exports power). In the neighborhood of τ_c the value U may be very large. Neither the significance of this phenomenon nor the conditions under which it comes about are fully understood. However it's resemblance to a Curie temperature is striking and suggests that further investigation is warranted.

Appendix B

Potential Value of Market Systems

Appendix B

Potential Value of Market Systems

We recognize the impact of a change in Q (regardless of whether it is caused by a change in P) by including the contribution of the change in system entropy to the total change in the value of the system. We do this by analogy to the thermodynamic identity:

$$
dU_P = \tau d\sigma + \mu dN - QdP \quad ;
$$
$$
dU_Q = \tau d\sigma + \mu dN + PdQ
$$
(B.1)

where τ is the value of system trading activity, σ is the system's entropy, N is the number of machines in the system, all of which were derived in Section 2.2, and μ is the system's migration potential as described by Sergeev (2003).

We can now consider the consequence of not permitting members of the system to choose their state (buy or sell). In most power distribution systems today, participants start out only as buyers, so $m = -\frac{1}{2}N$. The entropy of this system is

$$
\sigma = \log 2^{-N}
$$
(B.2)

which is the lowest possible entropy and so we will denote it σ_0. Any change in the state of any participant will necessarily increase the entropy to

$$
\sigma_1 = \log N 2^{-N} = \sigma_0 + \log N
$$
(B.3)

or

$$
\Delta\sigma = \log N .
$$
(B.4)

However for a participant to make such a change in state worth while, the system must return a gain in value ΔU to all N participants that is sufficiently large so that the player who made the change recover the cost of having made the change. The gain is

$$
\Delta U = 2qP
$$
(B.5)

where q is the quantity of power consumed by the participant. The value of this activity is given by

$$
\tau \approx \frac{\Delta U}{\Delta\sigma} = \frac{2qP}{\log N} .
$$
(B.6)

The problem we face is that for very large systems, $N \gg 1$ and $\tau \to 0$. As a result it takes an exceptionally high price P to stimulate any player to make the necessary structural change to permit a

change state. Therefore the only solution is to reduce the size of the system so that N is small enough to permit the entropy change to be large enough to realize a practically non-zero value for τ.

   Another way to think about this is to recall that the market provides the "hill climbing" mechanism as entropy maximizes. But when $m = \pm\frac{1}{2}N$, the value of $d\sigma/dm$ is essentially zero, making it impossible for the system to "find" a hill to climb.