6.3 Clustermodelle

6.3.1 Einleitung

Eine gängige Methode zur Untersuchung und Quantifizierung komplexer Strukturen, wie der des Konfigurationsraums, ist das Konzept des „random walks". Mit Hilfe plausibler Annahmen für diese Strukturen wurde z.B. in [39] erfolgreich versucht, das Relaxationsverhalten von Spingläsern zu simulieren. Unter Ausnutzung der Kenntnis der detaillierten Struktur der energetisch tiefliegenden Anregungen im Konfigurationsraum lassen sich jedoch vermutlich bessere Ansätze für derartige Simulationen finden. Ziel ist dabei stets die sinnvolle Vergröberung der gefundenen mikroskopischen Strukturen, um Aussagen über das makroskopische Langzeitverhalten des Systems zu gewinnen.

In diesem Abschnitt soll eine derartige Vergröberung in Form eines einfachen Clustermodells, welchem als Grundelement mikrokanonische Cluster zugrundeliegen, vorgeschlagen werden. Dabei wird zunächst kurz skizziert, wie das Problem des „random walks" im Zustandsraum unter Verwendung der Übergangsmatrix umgeschrieben werden kann in eine Mastergleichung (siehe z.B. [91, 92]).

Das Langzeitverhalten des Systems wird wesentlich vom Eigenwertspektrum der Übergangsmatrix im Bereich ihres betragsgrößten Eigenwerts bestimmt. Die exakte numerische Bestimmung der betreffende Eigenwerte läuft auf die Behandlung einer extrem großen allerdings nur sehr schwach besetzten Matrix, die alle mikroskopischen Feinheiten des Systems berücksichtigt, hinaus. Das Langzeitverhalten des Systems sollte jedoch nicht von allen diesen mikroskopischen Details der Struktur abhängen. Damit erscheint es sinnvoll, die Übergangsmatrix unter Ausnutzung der Kenntnisse über die wesentlichen Eigenschaften der Struktur des Konfigurationsraums zu vereinfachen. Ein gutes Modell sollte dabei zumindest das Verhalten der betragsgrößten Eigenwerte der ursprünglichen Übergangsmatrix reproduzieren. Zwei einfache Clustermodelle zur Vergröberung der Übergangsmatrizen werden im folgenden vorgestellt.

6.3.2 Random Walks im Konfigurationsraum

Angenommen ein System befindet sich mit der Wahrscheinlichkeit $p_i(t)$ zur Zeit t im Zustand i, dann bilden diese Wahrscheinlichkeiten den Vektor $\mathbf{p}(t)$ der Aufenthaltswahrscheinlichkeiten des Systems. Ein Zeitschritt eines „random walk" läßt sich dann in der Form

$$\mathbf{p}(t+1) = \mathbf{A}\mathbf{p}(t) \tag{6.4}$$

schreiben, wobei $\mathbf{A}$ die Übergangsmatrix bezeichnet, welche bei einer Gesamtzahl von M betrachteteten Spinkonfigurationen durch

$$A_{ij} = \begin{cases} w_{ij} & \text{für } i \neq j \\ 1 - \sum_{k=1,k\neq i}^{M} w_{ki} & \text{für } i = j \end{cases} \tag{6.5}$$

gegeben ist. Die Elemente w_{ij} entsprechen der Übergangswahrscheinlichkeit vom Zustand j in den Zustand i pro Zeitschritt. Diese Wahrscheinlichkeiten sind unter der Annahme von Einzelspinflips gegeben durch

$$w_{ij} = \begin{cases} 0 & \text{falls } i \text{ und } j \text{ nicht benachbart} \\ \frac{1}{N} & \text{falls } i \text{ und } j \text{ benachbart und } E_i \leq E_j \\ \frac{1}{N}\exp\left\{-\beta\left(E_i - E_j\right)\right\} & \text{falls } i \text{ und } j \text{ benachbart und } E_i \geq E_j \, , \end{cases} \tag{6.6}$$

wobei $\beta = 1/k_B T$ den Boltzmannfaktor bezeichnet. E_i bezeichnet die Energie des Zustands i und N entspricht der Maximalzahl der nächsten Nachbarn im Konfigurationsraum, welche direkt durch die Zahl der Spins im System gegeben ist. Man kann sich leicht davon überzeugen, daß die Gesamtwahrscheinlichkeit im System aufgrund der Definition der Übergangsmatrix **A** bei jedem Zeitschritt erhalten bleibt.

Die Matrix **A** ist nur im Spezialfall $\beta = 0$ symmetrisch. Im allgemeineren Fall $\beta > 0$ läßt sich **A** jedoch unter der Annahme des sogenannten detailed balancing in der Gleichgewichtsverteilung [91] durch die einfache Transformation

$$\begin{aligned} p_i' &= p_i \exp\left\{\beta E_i/2\right\} \\ A_{ij}' &= A_{ij} \exp\left\{\beta\left(E_i - E_j\right)/2\right\} \end{aligned} \tag{6.7}$$

in eine symmetrische Matrix überführen. Im folgenden sei daher **A** reell und symmetrisch.

Seien die Eigenwerte und -vektoren der Matrix durch

$$\mathbf{A}\mathbf{b}_i = \lambda_i \mathbf{b}_i \tag{6.8}$$

gegeben, dann sind alle Eigenwerte λ_i der Matrix **A** ebenfalls reellwertig. Es läßt sich weiterhin zeigen, daß aufgrund der Definition von **A** für alle Eigenwerte $|\lambda_i| \leq 1$ gilt (siehe auch [91, 92, 93]). Wenn die Eigenwerte entsprechend ihres Betrags sortiert sind, so ist der größte Eigenwert $\lambda_0 = 1$ trivial und der zugehörige Eigenvektor $\mathbf{b}_0$ beschreibt den Gleichgewichtszustand des Systems. Sofern der Zustandsraum zusammenhängend ist, sollte dieser triviale Eigenwert nur einfach entartet sein. Beim Auftreten getrennter unabhängiger Komponenten im Zustandsraum tritt eine entsprechende mehrfache Entartung auf und das Problem läßt sich in unabhängige Teilprobleme zerlegen.

Zur Berechnung der Zeitentwicklung des Systems erweist es sich als sinnvoll den Wahrscheinlichkeitsvektor zur Zeit $t = 0$ entsprechend des Eigensystems der Matrix **A** zu zerlegen:

$$\mathbf{p}(0) = \sum_{i=0}^{M-1} \alpha_i \mathbf{b}_i \, . \tag{6.9}$$

Damit läßt sich (6.4) umschreiben zu

$$\begin{aligned} \mathbf{p}(t) &= \mathbf{A}\mathbf{p}(t-1) \\ &= \mathbf{A}^t \mathbf{p}(0) \\ &= \mathbf{A}^t \sum_{i=0}^{M-1} \alpha_i \mathbf{b}_i \\ &= \sum_{i=0}^{M-1} \lambda_i^t \alpha_i \mathbf{b}_i \\ &= \sum_{i=0}^{M-1} e^{t \ln \lambda_i} \alpha_i \mathbf{b}_i \, . \end{aligned} \tag{6.10}$$

Demnach kann man jedem Eigenwert λ_i der Übergangsmatrix eine Relaxationszeit τ_i über

$$\frac{1}{\tau_i} = \ln \lambda_i \tag{6.11}$$

zuordnen, so daß eine ganze Schar von verschiedenen Relaxationszeiten im System existiert. Da alle Eigenwerte dem Betrag nach kleiner als Eins sind, tragen offenbar nur die Terme mit den größten Eigenwerten zum Langzeitverhalten des Systems bei.

Bei endlichen Systemen wird somit stets der betragsgrößte Eigenwert dominant und zu einem exponentiellen Einlaufen in den Gleichgewichtszustand führen. Bei unendlichen Systemen wird das Langzeitverhalten genau genommen vom Eigenwertspektrum der Übergangsmatrix in der Nähe des trivialen Eigenwerts Eins bestimmt, welches im Prinzip der Laplace-Transformierten des Zeitverhaltens entspricht [93, 94].

6.3.3 Eigenwerte der Übergangsmatrix

Mit der exakten Kenntnis der niederenergetischen Anregungen eines Systems läßt sich auch seine Übergangsmatrix berechnen. Ausgehend davon können dann numerisch die Eigenwerte der Matrix für verschiedene Temperaturen bestimmt werden. Die Abbildung 6.6 zeigt die Ergebnisse für die größten nicht trivialen Eigenwerte des dominanten Tals im Konfigurationsraum bezüglich der zweiten Anregung des oben eingeführten repräsentativen Systems. Die Bestimmung der Eigenwerte wurde mit Hilfe des sogenannten Lanczos-Algorithmus durchgeführt [95].

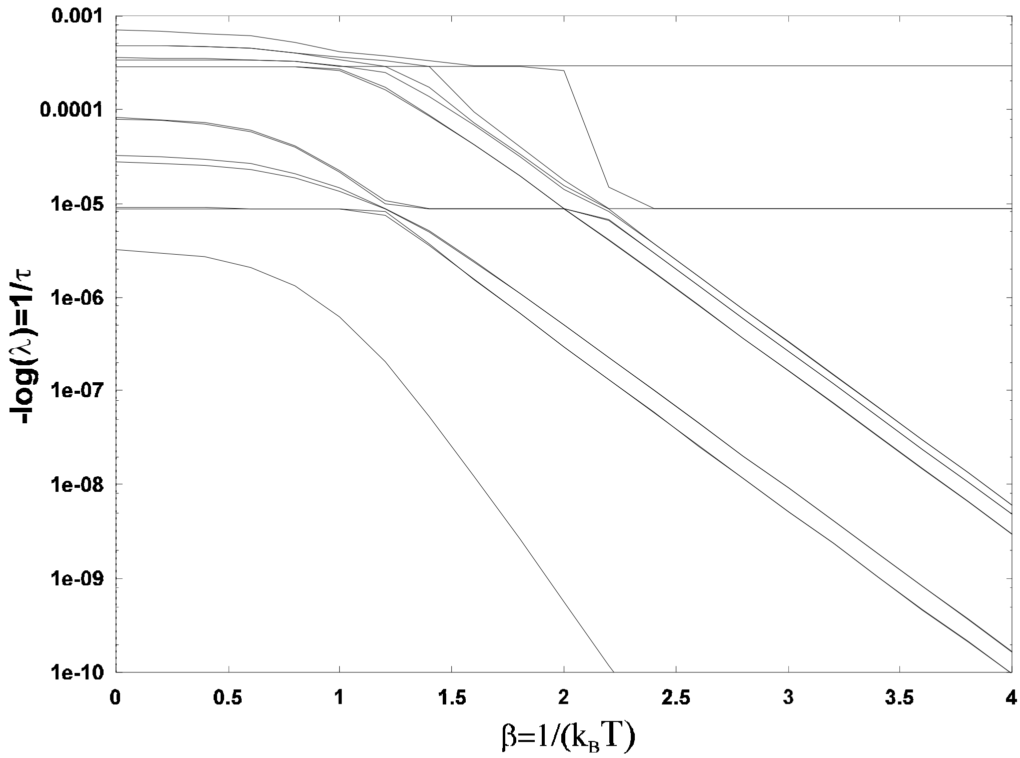

Abbildung 6.6 Betragsgrößte nicht triviale Eigenwerte der Übergangsmatrix eines repräsentativen Systems unter Berücksichtigung aller Zustände des dominanten Konfigurationsraumtals bezüglich der zweiten Anregung.

Es zeigen sich deutlich zwei unterschiedliche Typen von Eigenwerten. Einige der Eigenwerte führen zu divergierenden Relaxationszeiten bei abnehmender Temperatur. Berücksichtigt man, daß bis auf den größten alle Eigenwerte doppelt entartet sind, so existiert mit dem trivialen Eigenwert λ_0 für jedes der zwölf metastabilen Täler der Struktur in Abbildung 6.3 eine divergierende Relaxationszeit im System. Dies ist verständlich, da diese Täler bei verschwindender Temperatur getrennte Komponenten des Konfigurationsraums bilden und somit zu einer zusätzlichen Entartung des trivialen Eigenwerts führen. Die zugehörigen Eigenvektoren beschreiben demnach das Überqueren der energetischen Barrieren zwischen diesen metastabilen Tälern.

Die restlichen Eigenwerte sind im wesentlichen nur schwach von der Temperatur abhängig. Sie charakterisieren demzufolge die Relaxationsprozesse innerhalb der verschiedenen Täler, bei denen keine Energiebarrieren zu überwinden sind.

6.3.4 Einfache Clustermodelle

Die wohl einfachste Näherung zur Beschreibung der Struktur des Konfigurationsraums ist die Annahme eines völlig unstrukturierten Systems. In einem solchen einfachen Modell existieren keine Energiebarrieren. Daher sind alle Eigenwerte unabhängig von der Temperatur. Doch selbst bei unendlicher Temperatur ($\beta = 0$), bei welcher der Phasenraum als flach angesehen werden kann, treten sehr starke quantitaive Abweichungen zwischen dieser Näherung und den tatsächlich ermittelten Werten auf (siehe Anhang C). Offenbar sind nicht nur die energetischen Strukturen, die durch die Energiebarrieren gegeben sind, sondern auch die entropischen Barrieren, welche durch die innere Struktur der Täler bestimmt werden, wichtig für das Verständnis der Relaxationsprozesse im System.

Ein grundlegendes Element der gefundenen Energielandschaft sind mikrokanonische Cluster aus Zuständen mit jeweils der gleichen Energie. Die Relaxation innerhalb dieser Cluster hängt demzufolge nicht von der Temperatur ab. Die einzigen temperaturabhängigen Übergangswahrscheinlichkeiten beschreiben die Übergänge zwischen verschiedenen Clustern unterschiedlicher Energie. Daher sollte ein gutes Modell des Phasenraums zumindest die Übergänge zwischen verschiedenen Clustern und ein mehr oder weniger grobes Bild der inneren Struktur der einzelnen Cluster berücksichtigen.

Das gröbste Bild eines Einzelclusters besteht sicher in einer Vernachlässigung seiner inneren Struktur. Dabei würde dieser Cluster einfach auf einen einzelnen Gesamtzustand reduziert werden. Abbildung 6.7 zeigt die Struktur eines derartigen Modells (im weiteren als Modell A bezeichnet) für ein System aus drei mikrokanonischen Clustern, welche eine Energiebarriere bilden. Dabei bezeichnet n_i die Zahl der im Knoten i zusammengefaßten Zustände und p_i die mittlere Aufenthaltswahrscheinlichkeit im Knoten i.

Die Übergangsmatrix einer solchen einfachen Struktur ergibt sich zu

$$\mathbf{A} = \begin{pmatrix} 1-\frac{w_{31}}{n_1} & 0 & \frac{w_{13}}{n_1} \\ 0 & 1-\frac{w_{32}}{n_2} & \frac{w_{32}}{n_2} \\ \frac{w_{31}}{n_3} & \frac{w_{32}}{n_3} & 1-\frac{w_{23}}{n_3}-\frac{w_{13}}{n_3} \end{pmatrix}. \tag{6.12}$$

Für die Eigenwerte dieser Matrix erhält man

$$\begin{aligned} \lambda_0 &= 1 \\ \lambda_{1,2} &= 1-\frac{1}{2}\left(\frac{w_{13}}{n_3}+\frac{w_{23}}{n_3}+\frac{w_{31}}{n_1}+\frac{w_{32}}{n_2}\right) \\ &\quad \pm\frac{1}{2}\sqrt{\left(\frac{w_{13}}{n_3}-\frac{w_{23}}{n_3}+\frac{w_{31}}{n_1}-\frac{w_{32}}{n_2}\right)^2+\frac{4w_{13}w_{23}}{n_3^{\;2}}}\,. \end{aligned} \tag{6.13}$$

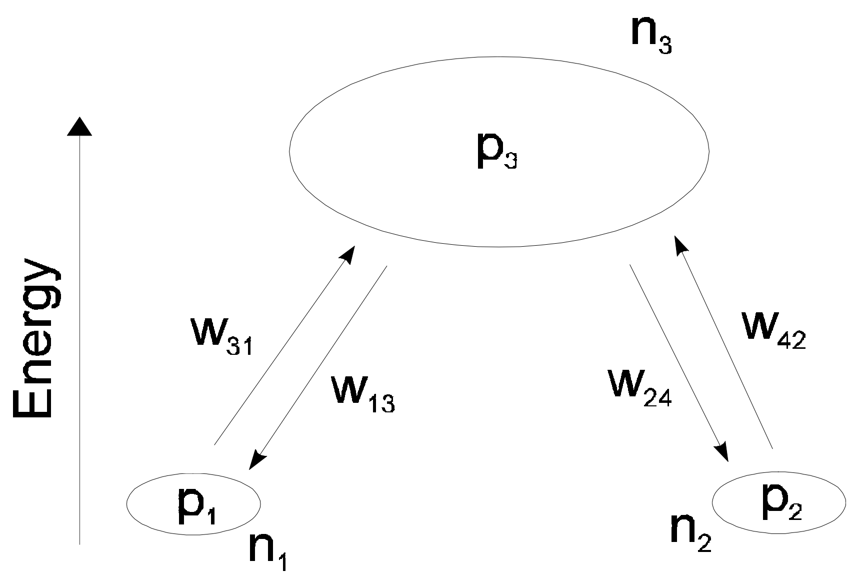

Abbildung 6.7 Schematisches Bild des Clustermodells A.

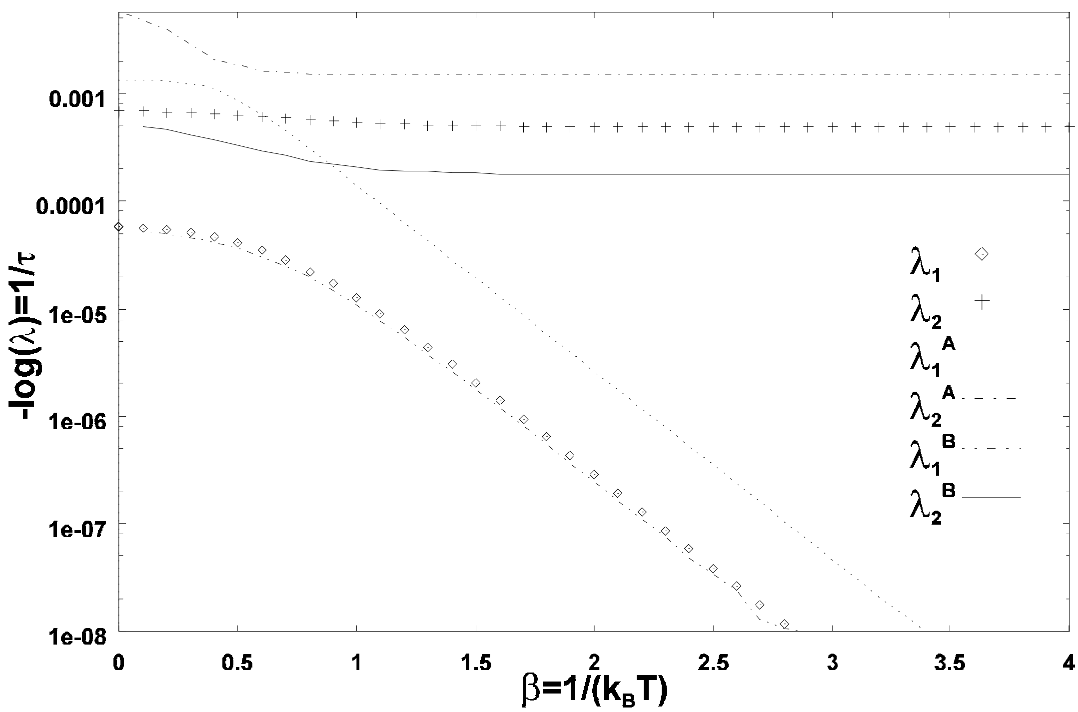

Abbildung 6.8 Vergleich der betragsgrößten nicht-trivialen Eigenwerte λ_1 und λ_2 für die verschiedenen Clustermodelle.

Die temperaturabhängigen Übergangswahrscheinlichkeiten zwischen den verschiedenen Clustern können durch einfaches Abzählen der Verbindungen zwischen diesen Clustern und der Zustände in den Randregionen zum jeweils anderen Cluster ermittelt werden.

Um das Modell zu testen, wurde eine typische Untermenge von drei großen mikrokanonischen Clustern aus der gefundenen Struktur des repräsentativen Systems isoliert und die Eigenwerte der Modellmatrix (6.12) mit denen der Gesamtmatrix verglichen (Abb. 6.8).

Das qualitative Verhalten der Eigenwerte des Modells A stimmt gut mit den direkt ermittelten Eigenwerten überein. Quantitativ treten jedoch starke Abweichungen auf. Der Relaxationsprozeß im Modell ist wesentlich zu schnell. Die Ursache dafür liegt vermutlich im unterschätzten Einfluß der Relaxation innerhalb der verschiedenen Cluster.

Daher wurde in einem verbesserten Modell B versucht, die innere Struktur der Cluster zu berücksichtigen. Für tiefe Temperaturen sollten die wahrscheinlichsten Zustände eines Clusters in der Nähe der Randregionen zu energetisch tiefer liegenderen Clustern zu finden sein. Für hohe Temperaturen sollte dann das Innere des Clusters wichtiger werden. Deshalb wurde in Modell B ein Cluster in seine Randregionen und einen inneren Anteil zerlegt.

Wenn $\mathbf{p}(t)$ den Gesamtvektor der Aufenthaltswahrscheinlichkeiten eines mikrokanonischen Clusters darstellt, dann läßt sich die Gesamtwahrscheinlichkeit einer Randregion r entsprechend

$$p_r(t) = \mathbf{p}(t)\mathbf{r} \tag{6.14}$$

berechnen, wobei die Komponenten des Randvektors $\mathbf{r}$ für Spinkonfigurationen, die zum betreffenden Rand gehören Eins und ansonsten Null sind. Damit die Wahrscheinlichkeit im Gesamtcluster erhalten bleibt, wird die noch fehlende innere Region bei k Randregionen entsprechend

$$\mathbf{r}_0 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} - \sum_{i=1}^{k} \mathbf{r}_i \tag{6.15}$$

gewählt. Da die Komponenten von $\mathbf{r}_0$ bei überlappenden Rändern auch negativ werden können, stellt diese innere Region nur einen fiktives Modell für den von den Randregionen nicht erfaßten Teil des Gesamtclusters dar.

Approximiert man den Vektor $\mathbf{p}(t)$ durch seine Zerlegung bezüglich der größten Eigenwerte der Übergangsmatrix des betreffenden mikrokanonischen Clusters entsprechend

$$\mathbf{p}(t) = \sum_{i=0}^{k} a_i(t)\mathbf{b}_i\,, \tag{6.16}$$

so erhält man

$$p_r(t) = \sum_{i=0}^{k} a_i(t)\mathbf{b}_i\mathbf{r}\,. \tag{6.17}$$

Nach einem Zeitschritt ergibt sich dann offenbar

$$p_r(t+1) = \sum_{i=0}^{k} a_i(t)\lambda_i \mathbf{b}_i \mathbf{r} \,. \tag{6.18}$$

Für m Randregionen eines Gesamtclusters ergeben sich $2m$ unabhängige Gleichungen vom Typ (6.17) und (6.18) aus denen man durch Eliminierung der $k+1$ Unbekannten $a_i(t)$ die reduzierte Übergangsmatrix zwischen den m Randclustern gewinnen kann, falls $m = k+1$. Somit benötigt man zur Approximation eines mikrokanonischen Clusters mit m Randregionen wenigstens die $m-1$ betragsgrößten Eigenwerte und Eigenvektoren des betreffenden Clusters.

Da die Cluster höherer Anregungen im allgemeinen wesentlich größer sind als niederenergetische Cluster, wurde zunächst nur der obere der drei Cluster des Testsystems in verschiedenen Regionen aufgeteilt, so daß sich für Modell B eine Struktur entsprechend Abbildung 6.9 ergibt.

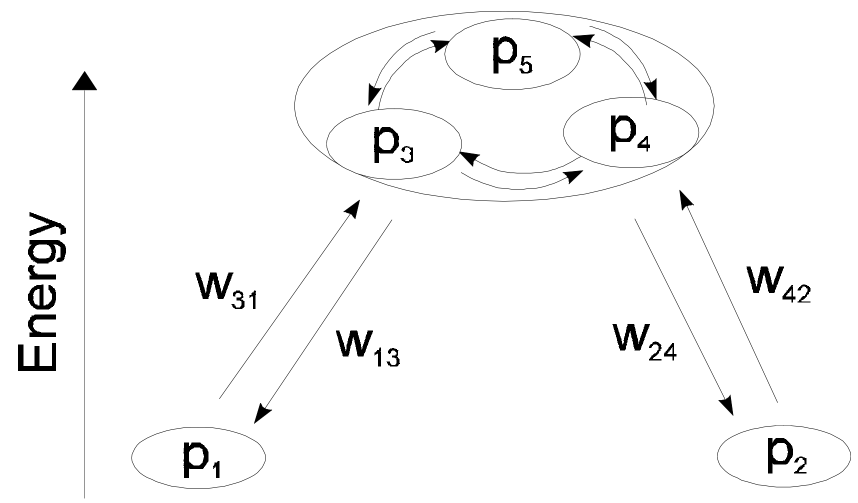

Abbildung 6.9 Schematisches Bild des Clustermodells B.

Der entsprechend diesem Modell ermittelte betragsgrößte nicht triviale Eigenwert λ_1 stimmt im gesamten Temperaturbreich sowohl qualitativ als auch quantitativ mit dem direkt ermittelten Eigenwert überein. Die Näherung für λ_2 stimmt zumindest qualitativ gut mit dem exakten Wert überein.

Eine Überprüfung der Güte der vorgestellten Modelle bei ihrer Anwendung auf größere Systeme mit wesentlich mehr mikrokanonischen Clustern steht noch aus. Trotzdem scheint es prinzipiell möglich zu sein, durch eine gezielte Vergröberung der exakten mikroskopischen Energielandschaft eines Systems die zugehörige komplexe und oftmals numerisch nur schwer zugängliche Übergangsmatrix drastisch zu reduzieren, wobei der Informationsverlust für die das Langzeitverhalten bestimmenden betragsgrößten Eigenwerte nur gering ist.