<u>Trellis-Coded-Modulation (TCM)</u>
<u>Discovered (Invented) by G. Ungerboeck</u>

"Channel Coding with Multilevel/Phase Signals"

- IEEE Trans. Info. Theory, Vol. IT 28, No. 1, Jan 1982, pp 55-67.

⇒ Regard channel coding and modulation as an entity together

- as proposed first by <u>Massey</u>- as opposed to Coding and Modulation separately.

<u>Classical View</u>

Error Control Coding ⇒ Increase in spectrum bandwidth
Decrease in SNR

Due to the fact that the rate of the encoder o/p is greater than the rate of the encoder i/p

- <u>assuming identical alphabets</u> – by 1/R
  (Consider R = ½ convolutional encoder)

⇒ Coding could only provide gains in energy efficiency when the possibility of bandwidth expansion is present.

<u>Ungerboeck</u>

-argued that the modulation symbol set could be enlarged when coding is used relative to that needed for "uncoded" case.

If the signal set dimensionality per info bit is unchanged the power spectrum remains unchanged. (<u>no BW expansion</u>)

* <u>The signaling rate does not change</u>.

i.e., Coding and modulation is performed with respect to Euclidean distance between coded modulation o/p sequences.

⇒ MLD soft decision decoding.

<u>Set Partitioning</u>

Consider a constellation having M points in an N dimensional space. (1, 2 dim. originally)

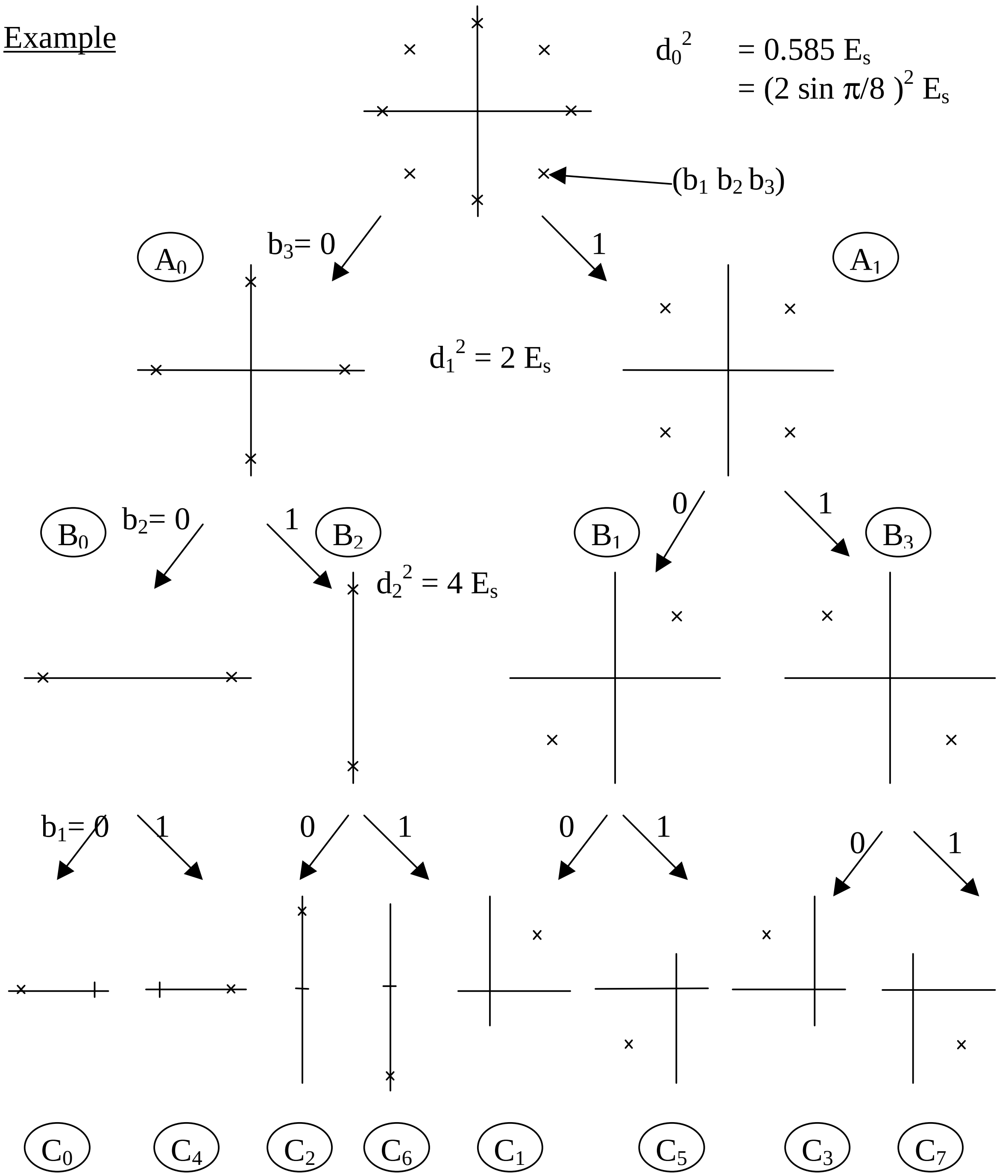
<u>Example</u>

M-ary PAM, M-ary QAM, PSK or subsets of N-dimensional lattices.

- First partition the original set into p_1 equal sized disjoint subset or cosets. ⇒ Each of size M/p_1

- $p_1 \Rightarrow A_1, A_2, \ldots, A_{p_1}$ subsets. So that within each subset the minimum ED between signal points is maximal.
- Next each of these subsets are further split into p_2 subsets denoted by B_0, B_1, ...
- Proceeding recursively original constellation may be eventually decomposed into single point subsets.
- Typically, the splitting factors p_i are equal $p_i = 2$ usually (not essential.)

Example

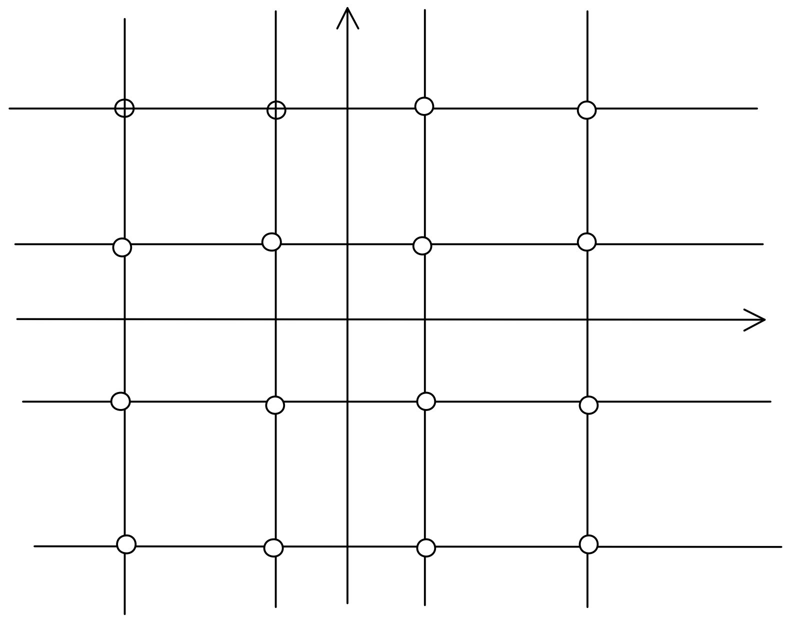


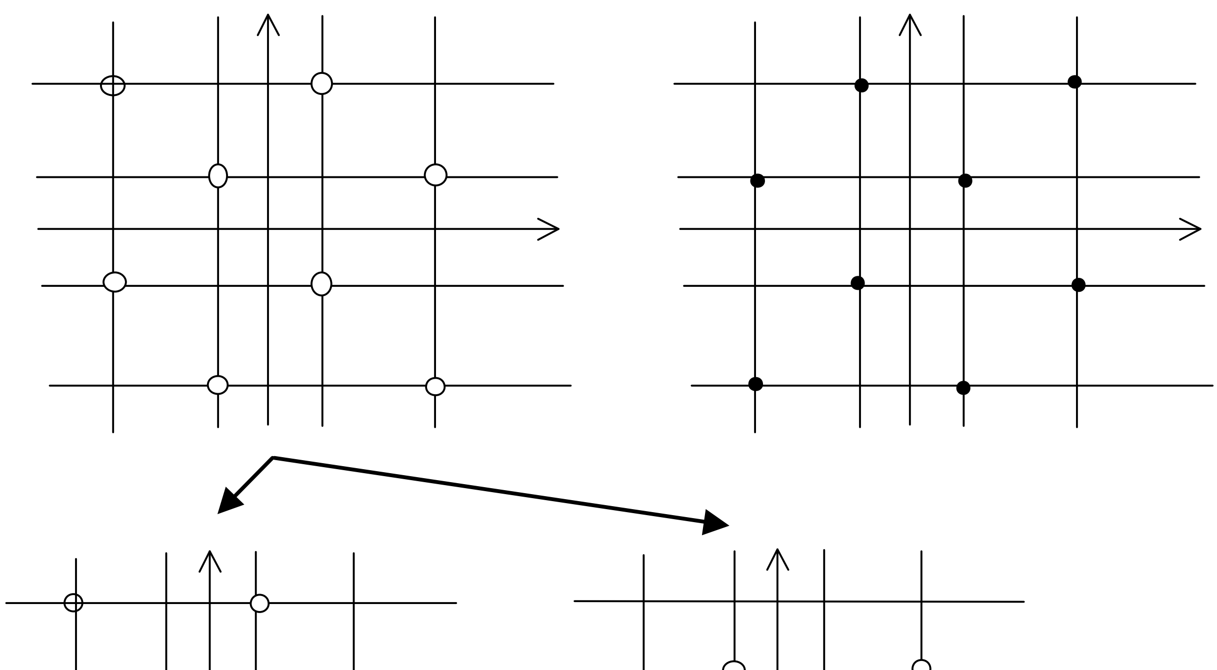
Set partitioning of 8-PSK with associated bit labeling

Example 2

Partitioning of the 2-dimensional Lattice Z^2. (2a is the constellation spacing along each signal-space axis.) – 16 QAM, 32-cross.

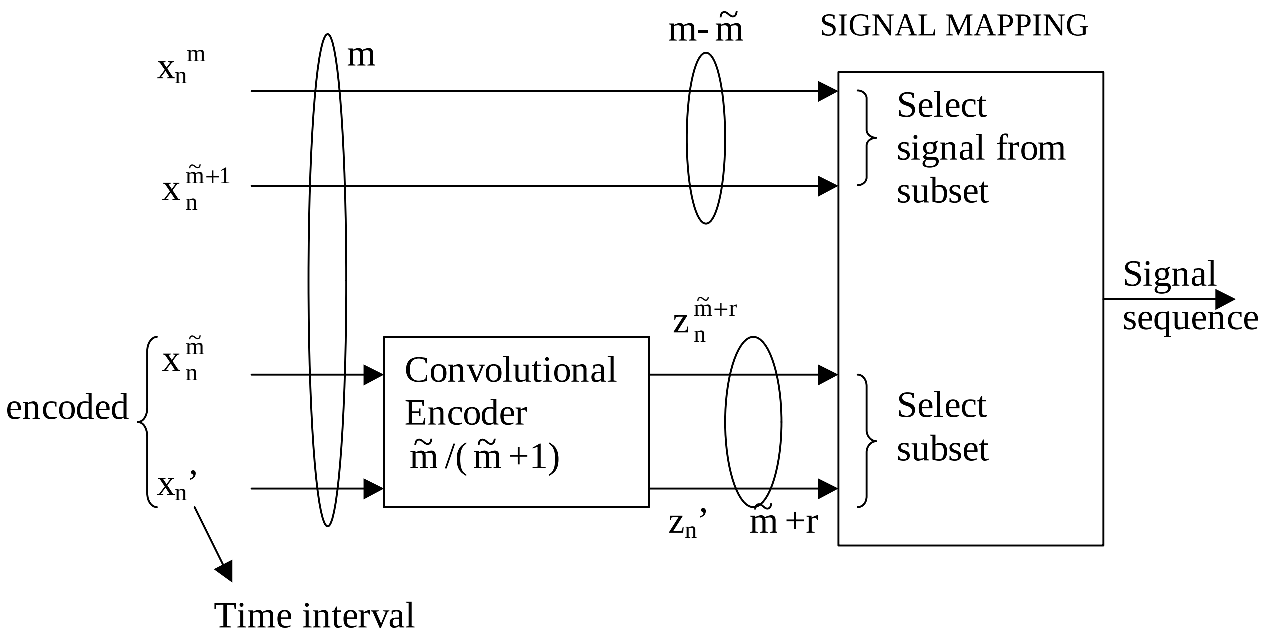
16 QAM Signal Constellation- Set Partitioning





*Trellis Coding for band limited channels

1) If m bits to be transmitted/modulation interval
2) $\tilde{m} \leq m$ bits expanded into $\tilde{m}+1$ ($\tilde{m}+r$) coded bits by $\tilde{m}/(\tilde{m}+1)$ convolution encoder
3) Use these $\tilde{m}+1$ bits to select $2^{\tilde{m}+1}$ subsets of a 2^{m+1} signal set.
4) Remaining $m-\tilde{m}$ uncoded bits determine which of the $2^{m-\tilde{m}}$ signals in the subset is transmitted.



Modulation has 2^{m+r} points in N dimensions. (r = 1 usually)

Hand Design of Codes

(*) Selection of the trellis size – No. of states

$$S = 2^v, \qquad v \text{ – memory}$$

(*) m bits/trellis level.

$\Rightarrow$ 2^m branches leaving and merging with each state.

$$m = 3 \Rightarrow 2^3 = 8$$

This can be divided as

i) eight single branches

$$\tilde{m}=3$$

ii) four groups of two

$$\tilde{m}=2$$

iii) two groups of 4

$$\tilde{m}=1$$

which one is best is not obvious.

$\Rightarrow$ We assign constellation subsets of proper size to various state transitions.

Ungerboeck's Rules

1) Employ all subsets equally often m labeling the trellis.
2) For subset assignments that share a common splitting state or merging state, choose subsets between which the minimum "intersect" distance is largest.

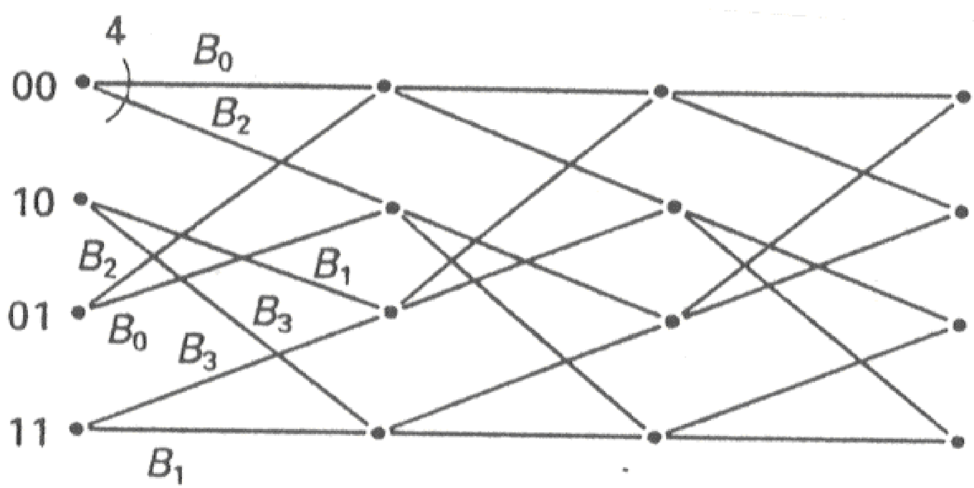
Example

4 state code for m = 2, 8-PSK.

Throughput is m =2 bits/modulator symbol.

No. of branches/state = $2^2 = 4$

Assign branching as 2 sets of 2 as shown below, i.e., $\tilde{m}=1 \Rightarrow 1$ bit left uncoded.



Trellis labeling for 4-state, 8-PSK code. Subsets have size 2.

Rule 1 There are 4 antipodal subsets. Trellis requires 8 subset assignments $\Rightarrow$ each subset used twice.

Rule 2 For branches leaving (merging) assign subsets that differ by 90° rotation.

(*) Presence of "parallel transitions" – TCM design only
Not in convolutional codes.

⇒ Single step error events are possible.

For ML decoding "free distance" – min ED between any two coded sequences, is the figures of merit (at high SNR)

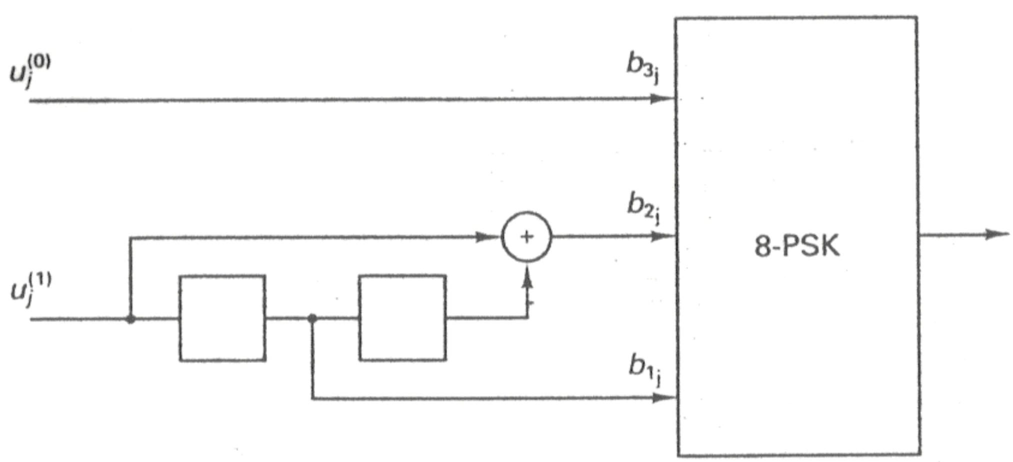
ED for parallel transition $d_1^2 = 4\ E_s$
3 stage error $d_3^2 = 2\ E_s + 0.585\ E_s + 2\ E_s$
$= 4.585\ E_s$

⇒ $d_f^2 = 4\ E_s$ (four state code 8-PSK)

Each code symbol carries the energy of 2.
info bits $E_s = 2\ E_b$
$d_f^2 = 8\ E_b$

Uncoded case
$d_{f_u}^2 = 2\ E_s = 4\ E_b$

$$\text{Coding gain} = 10\log_{10}\left\{\frac{d^2_{free(coded)} / d^2_{free(uncoded)}}{\text{Energy}_{coded} / \text{Energy}_{uncoded}}\right\}$$



Feedforward encoder realization of best 4-state code for 8-PSK.

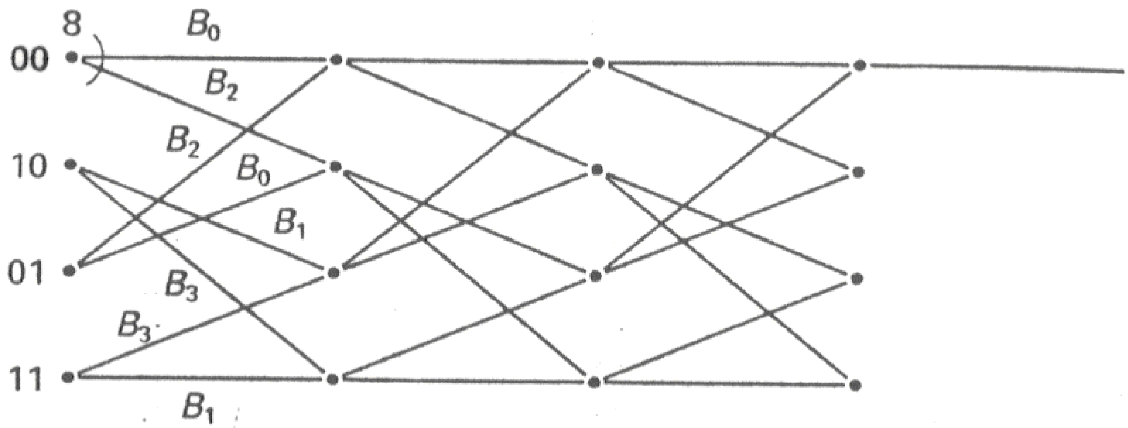
(*) For – four sets of one branching we can show that
$d_f^2 = d_2^2 = 2.585\ E_s$ – suboptimal trellis structure.
- No parallel transitions.

Note : This trellis structure is identical to the best binary Hamming distance code with R = 2/3. Even with best assignment of 8 PSK to o/p 3 tuples we produce an inferior code for AWGN channel.

Example

16 QAM – TCM with m =3 bits/symbol.

4 state trellis with $\tilde{m}$ =3 coded bit



Trellis labeling for 4-state, 16 QAM code. Note similarity with 4-state, 8-PSK design. Subsets have size 4.

- Single step error events are dominant.
  $d_f^2 = 16\ a^2$ where 2a is the along axis spacing in the constellation.

  E_s = Average energy per symbol = $10\ a^2$
  $\therefore\ d_f^2 = 1.6\ E_s = 1.6\ (3E_b) = 4.8\ E_b$

  3 info. bits

  For the uncoded 8-PSK
  $d_{f_u}^2 = 0.585\ E_s = 0.585\ (3E_b) = 1.755\ E_b$

  $\therefore$ Coding gain = 4.8/1.755 $\simeq$ 4.4 dB

<u>Referring to the Trellis encoder given the design questions are, for a given m and S ($=2^v$, v = no. of binary memory elements)</u>

1) What constellation should be used ?
2) What is the proper (or what form of trellis branching should be used)?
3) What is the best subset labeling pattern or, equivalently, the best encoder?

Rearranging the constellation issue

(*) Codes with constant energy per transmitted symbol may be important – M-PSK.

(*) Constellation Size : Ungerboeck argued that expansion by a factor 2 is adequate for 2-D constellations e.g. uncoded M QAM $\Rightarrow$ coded 2M – QAM.

$\Rightarrow$ dimensionality per info. bit remains unchanged

$\Rightarrow$ to first order, so is the power spectrum i.e Bandwidth

(*) 8 points constellations
for R = 2 bits/symbol
16 point $\Rightarrow$ 3 bits/symbol.

(*) Large complexity codes by computer search.
combining CC's with set partitioning, i.e., we can use the labeling implied by the set partioning powers and the search over the class of binary CC's with S states, determining the free distance for each such code.

(*) Use systematic encoders with o/p feedback.
(Any feedforward nonsystematic CC can be realized as one.) $\Rightarrow$ Noncatastrophic.

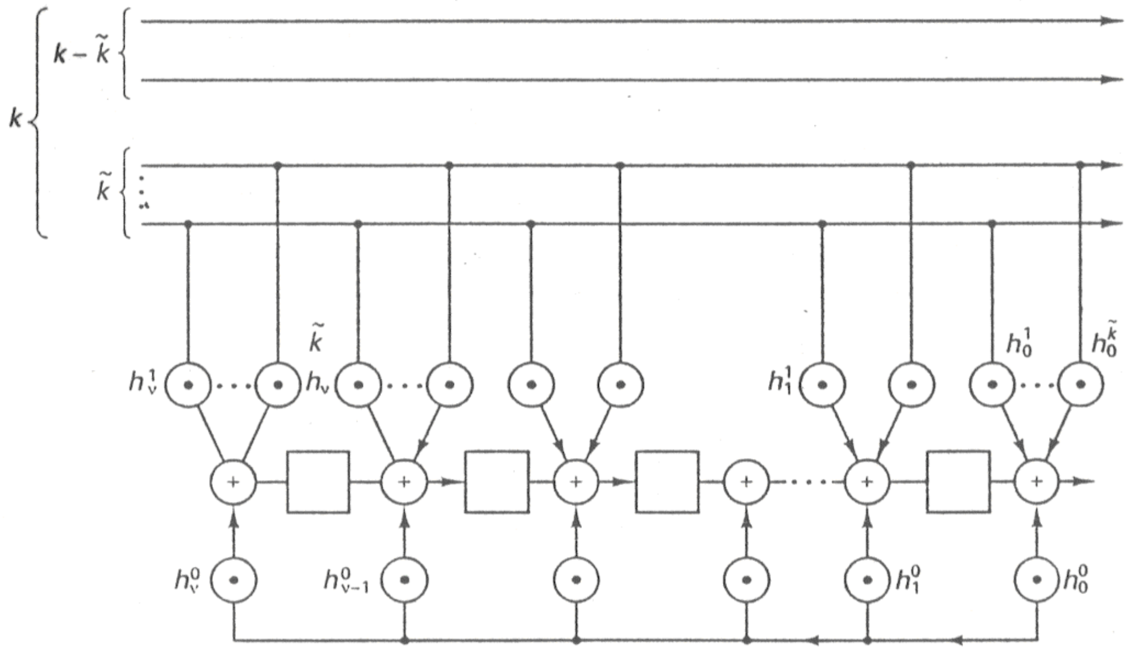
Systematic encoders with feedback are specified by listing the parity check polynomials.

$$h^i(D) = h_0^i + h_0^i D + ... + h_v^i D^v \text{ for } i = 0, 1, ..., m.$$

For encoders with only coded bits

$h^i(D) = 0$, $\tilde{m} < i < m$ which says that the uncoded bits need not obey any parity constraints.

$$m = k,\ \tilde{m} = \tilde{k}$$

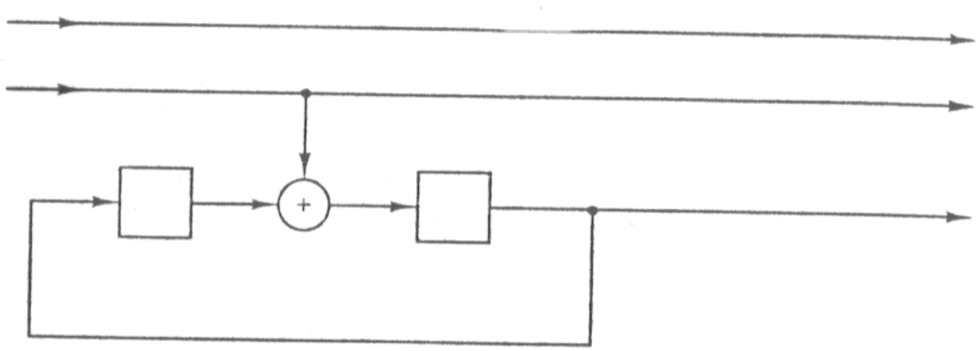


Generic 2^ν – state encoder in systematic form with feedback.

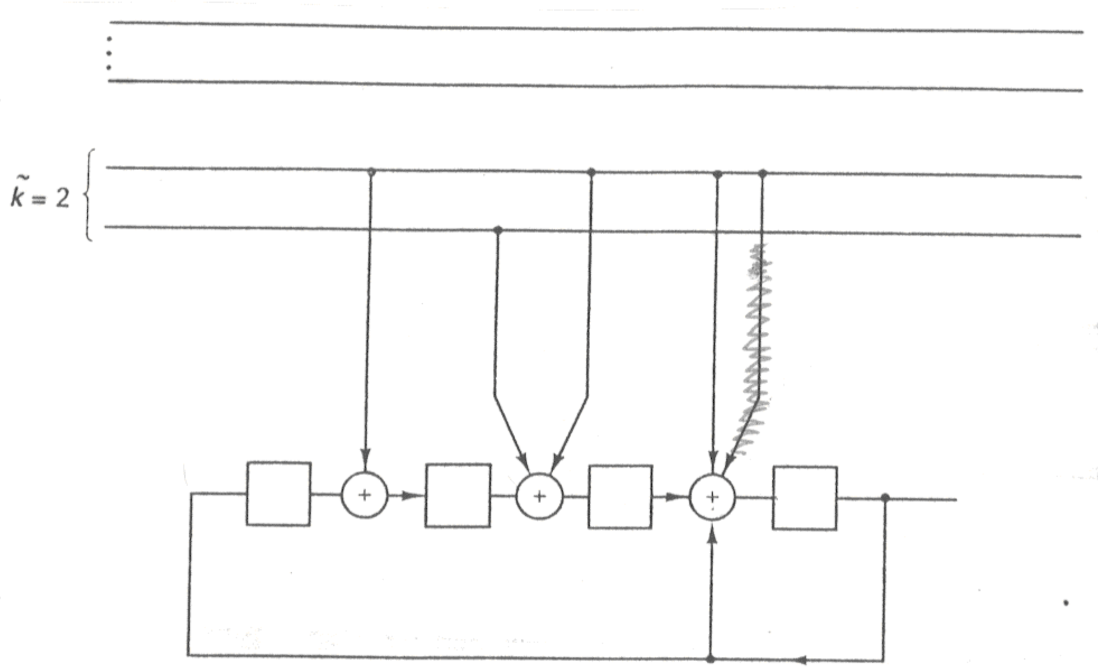
(*) The figure illustrates the case where only $\tilde{m}$- inputs (coded bits) influence the state vector.

$h^0(D)$ provides the feed back connections.
Usually $h^i(D)$ are specified in octal notation.

Example
For 8-PSK, 4-state encoder
$h^0(D) = 1 + D^2$; $h^1(D) = D$; $h^2(D) = 0$



Four-state encoder for 8-PSK in systematic form with feedback.
$h^0(D) = 1 + D^2$; $h^1(D) = D$; $h^2(D) = 0$



Sixteen-state encoder for QAM constellations. $h^0(D) = 23_8$;
$h^1 = 4_8$; $h^2 = 16_8$, other $h^i = 0$.

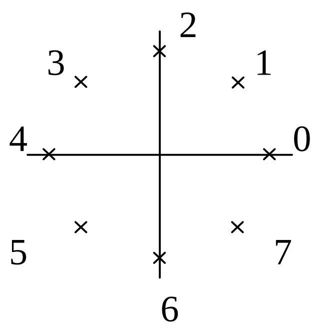
- Bigalieri has shown that for Ungerboeck-style TCM, i.e., set partitioned constellations proceeded by CC's, the power spectrum is in fact unaltered. (symmetries of the signaling process – mostly due to)
  "Ungerboeck Codes Do Not Shape Power Spectrum"
  IEEE Trans. Info Theory Vol. IT-32, pp-595-596, July 1986.

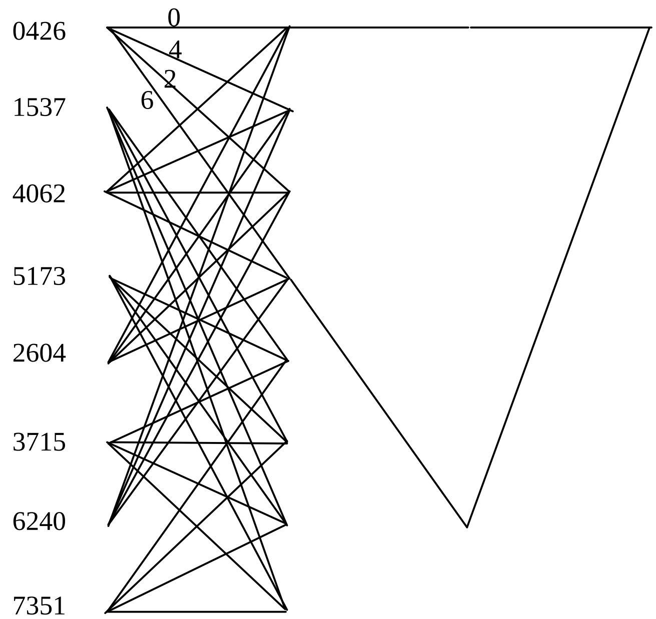
Example
Note that parallel transitions imply that single signal error events can occur. This limits achievable free Euclidean distance to the minimum distance in the subsets of signals assigned to parallel transitions.
On the other hand, parallel transitions reduce the "connectivity" in the trellis and allow extension of the minimum length of the multiple signal error events.

(*) 4 state – 8 PSK – with parallel transitions- best solution-3 dB gain

But with 8 and more states, only trellis structures with distinct transitions can be of interest, otherwise free ED gain would remain limited to 3 dB.

8 PSK





3.6 dB gain over 4-PSK

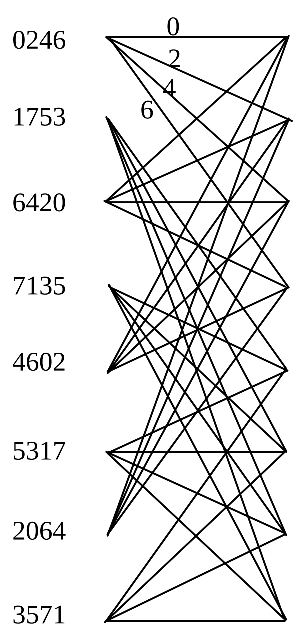
(*) Transitions originating from the same state or merging into same state assigned signals from either subset A_0 or A_1

Decode the received signal in 2 steps:

(a) Subset decoding – within the signal subset assigned to the parallel transitions, choose the signal closet to the received channel o/p. Store these signals together with their squared distances.

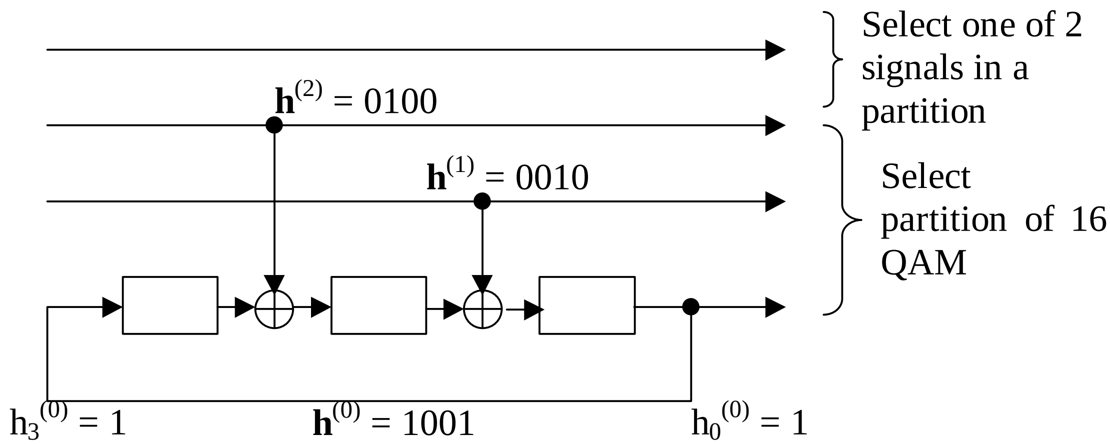
(b) Use the VA to find the signal path that is cloest to the received signal using these subset winners and their ML metrics.

Example
8-State, 16 QAM



Examples (TCM)
Encoder configurations:

			Octal		Binary
$\tilde{m}=2$	$\mathbf{h}^{(0)}=11$	$\rightarrow$	001001	$\rightarrow$	1001
$m=3$	$\mathbf{h}^{(1)}=02$	$\rightarrow$	000010	$\rightarrow$	0010
$v=3$	$\mathbf{h}^{(2)}=04$	$\rightarrow$	000100	$\rightarrow$	0100



It is shown that the number of coded bits, 2, remains unchanged for most of MPSK, MQAM schemes regardless of the memory and signal points. This is for best TCM codes in AWGN channel.
[Ref: Ungerboeck 1982, 1987.]

Decoding TCM over AWGN Channels

In the previous sections the minimum Euclidean distance *(d_{free})* was always used as a guideline to design good codes. Actually this result comes from the analytical upper bound on Bit Error Rate (BI/LR) of TCM for thc AWGN channels.

In general the TCM trellis assigns a sub-constellation of signals to each branch in the trellis. Viterbi decoding is performed m the two steps:

1. At each branch in the trellis the receiver compares the received signal to each signal allowed for that branch. The identity of the allowed signal closest to the received signal is saved in the memory and the branch is labeled with a metric proportional to the distance between the two signals.

2. The Viterbi Algorithm is then applied to the trellis with surviving partial paths corresponding to partial signal sequences that are closest to the received sequence. The maximum likelihood path is then the complete signal sequence closest in Euclidean distance to the received sequence.

When the Viterbi decoder is used, the performance analysis for TCM system is similar to that for convolutional codes. Following the derivation, the upper bound on the node error rate (i.e. the rate whenever a nonzero path leaves a given node and reemerges with the correct path at a later node) is given by:

$$P_e \le Q\left(\sqrt{\frac{d_{free}^2 E_s}{2N_0}}\right) . exp\left(\frac{d_{free}^2 E_s}{4N_0}\right) . T(D)\Big|_{D=exp(E_s/4N_0)} \quad (1)$$

And the upper bound on BER is:

$$P_b \le \frac{1}{m} . Q\left(\sqrt{\frac{d_{free}^2 E_s}{2N_0}}\right) . exp\left(\frac{d_{free}^2 E_s}{4N_0}\right) . \frac{\partial T(D,I)}{\partial I}\Big|_{D=exp(E_s/4N_0), I=1} \quad (2)$$

Where *T(D)* is the generating function of the directed graph which is related to the state diagram of the trellis code. *T(D,I)* is the augmented version of *T(D)* with components of the *I* terms denote the number of information bit errors associated with each error event. N_0 is the one-sided noise spectral density of AWGN, E_s. is the average energy of the signal constellation and *m* is the number of information bits carried by each symbol.

The derivation of generating function can be quite complicated and becomes intractable as the number of states in the TCM increases. However at high Signal-to-Noise Ratio (SNR) (1), (2) can be approximated quite accurately without the use of *T(D,I) :*

$$P_e \approx N(d_{free}) Q\left(\sqrt{\frac{d_{free}^2 E_s}{2N_0}}\right) \quad (4)$$

$$P_b \approx \frac{d_{free}}{m} Q\left(\sqrt{\frac{d_{free}^2 E_s}{2N_0}}\right) \quad (5)$$

Here $N(d_{free})$ is the average number of sequences that are distance d_{free} from the transmitted sequence and $b_{d_{free}}$ is the total number of the information bit errors associated with erroneous paths that are distance d_{free} from the transmitted path, averaged over all possible transmitted path. Equations (4) and (5) justify the importance of minimum Euclidean distance to the asymptotic coding gain.

Some Applications of Trellis-Coded Modulation in CCITT Recommendations

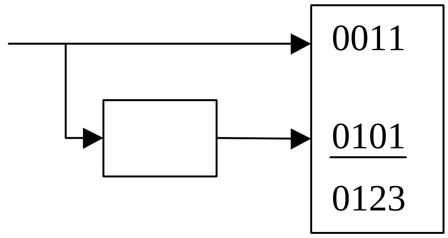
V.32 is intended for 9.6 Kbps traffic over two-wire telephone lines and 14.4 Kbps traffic over over-wire circuit. The TCM code is used in V.32 modem in order to obtain higher noise immunity over uncoded transmission at the same data rate, transmission power and bandwidth limit.

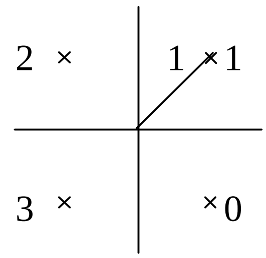
The V.32 TCM code uses m = 4 binary input symbols, signal set has $2^{(m+l)} = 32$ signal vectors in the 32-CROSS constellation. A convolulional encoder with k 2 input, n=3 output, and memory M = 3 is implemented. This is a non-linear encoder. With this implementation a asymptotic coding gain of 4dB is obtained.

The same nonlinear encoder was later adopted for use with 64-QAM and 128-CROSS in V. 17 (14.4 Kbps FAX traffic over standard phone lines) and V.33.

TCM technique is relatively young. Several questions concerning real coding gains, performance under channel impairments other than AWGN channels (especially for mobile radio channels) and implementation complexities are actively being studied.

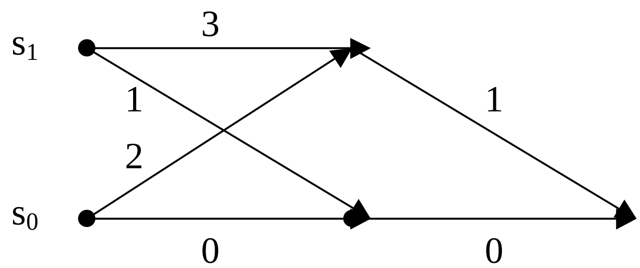
MTCM





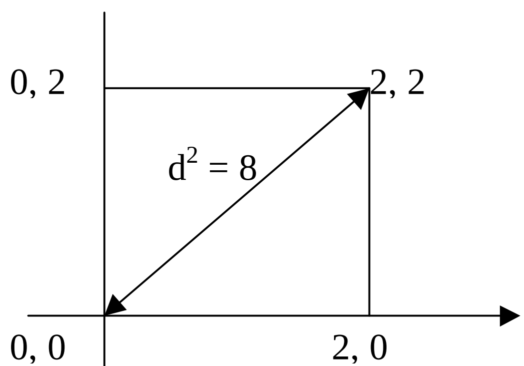
$d_{free}^2 = d^2(0, 2) + d^2(0, 1) = 4 + 2 = 6$

TCM Trellis:

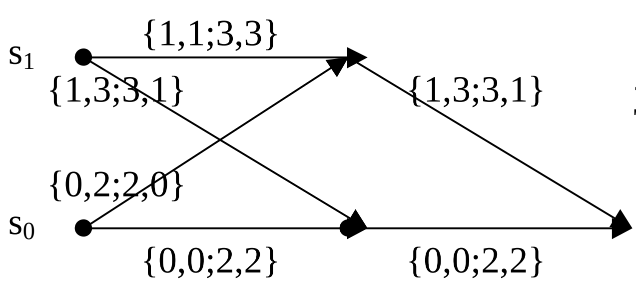


$s_0 \rightarrow \{0, 2\}$, $s_1 \rightarrow \{1, 3\}$

Now allow the encoder to select r = 2 signals (with replacement) from these respective partitions for each trellis branch.
Thus for $s_0 \rightarrow \{0,0 ; 0,2 ; 2,0 ; 2,2\}$
These can be viewed as the vertices of a square.



Since these are four possible two-signal sequences but only two successor states for s_0 – i.e s_0, s_1 there will be parallel transitions. The two signal sequences are partitioned into pairs such that the inter-sequence distance in each pair is maximized.



2 state MTCM
Trellis with r = 2

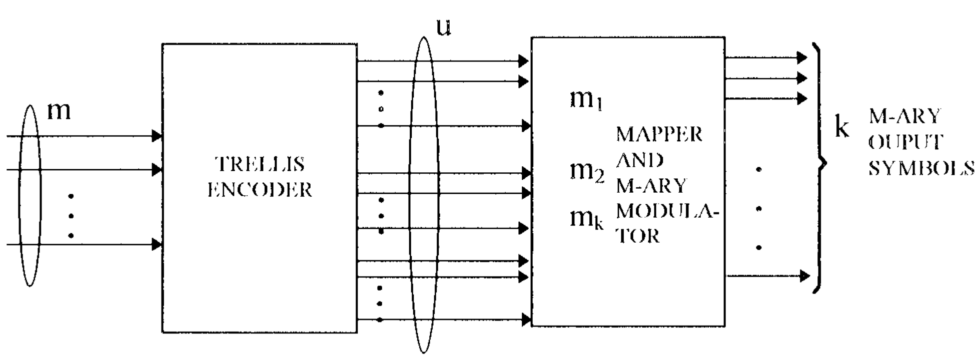
$d^2_{free,\ parallel} = d^2(0, 2) + d^2(0, 2) = 4 + 4 = 8$
$d^2_{free,\ nonparallel} = d^2(0, 0) + d^2(0, 2) + d^2(0,1) + d^2(0,3)$
$= 0 + 4 + 2 + 2$
$= 8$

Coding gain over TCM = 8/6 = 1.33 (1.25 dB)
Coding gain over BPSK = 8/4 = 2 (3.01 dB)

This further development of TCM technique was made by Divsalar and Simon where the authors demonstrated a trellis coded-modulation technique referred to as multiple trellis coded - modulation (MTCM). It was shown that MTCM can achieve a significant increase in coding gain relative to conventional TCM by increasing the number of signals transmitted per each trellis branch.

The principle behind MTCM is to design a rate *nk/(n+I)k* encoder and combine it (again through a suitable mapping) with 2^{n+1} - point signal constellation and to output k of these signal points in each transmission interval. Since in each transmission interval, /or bits enter the encoder and k symbols leave the modulator it is guaranteed to have unity bandwidth expansion relative to a 2^n - point uncoded system.

A multiple trellis encoder has m binary input bits and u binary output bits which are mapped into *k* M-ary symbols in each transmission symbol. Such a system has throughput *m/k bits/s/Hz* which has the same bandwidth with $2^{m/k}$ - point signal constellation. In the considered system the partition is carried out as follows: u binary encoder output bits are partitioned into k group of $m = \log_2 M$ bits each. Each of these groups results in an MPSK output symbol. Clearly, to achieve this result the parameters *u, k, M* have to be chosen so that $u = k \cdot \log_2 M$ (here $m_1=m_2=\ldots=m_k$).

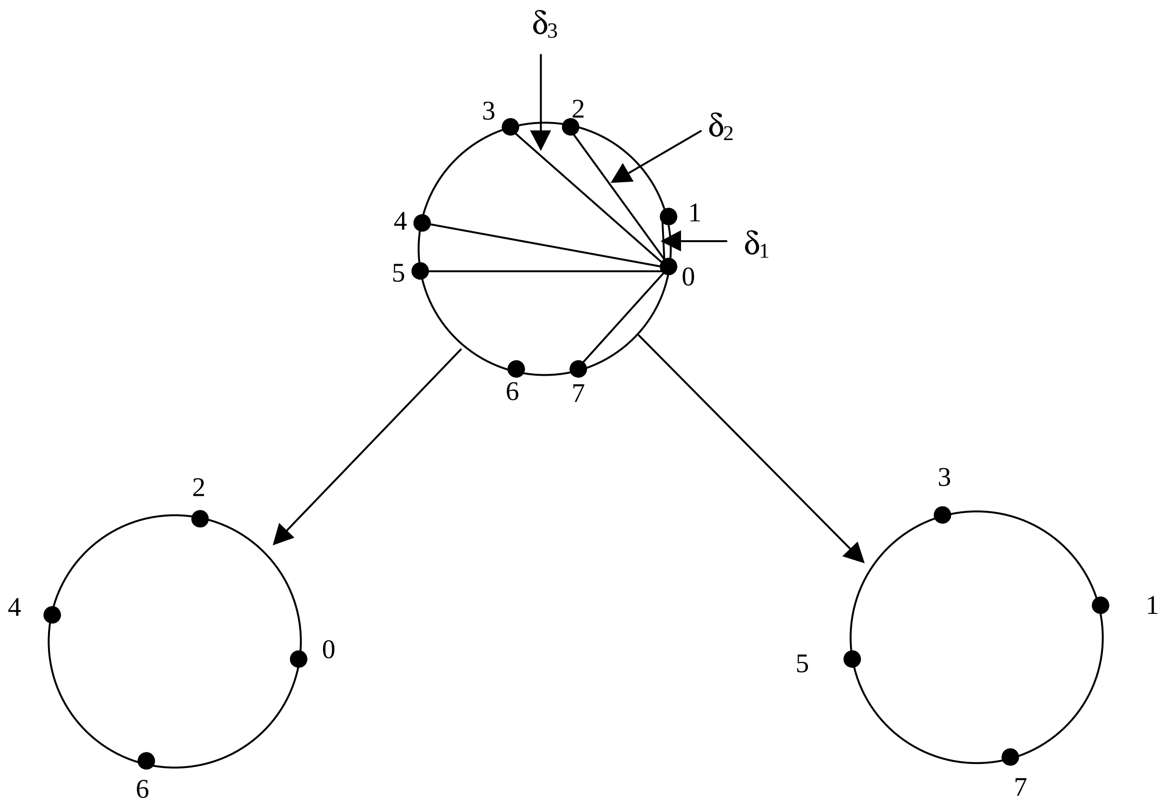


Multiple Trellis Encoded M PSK Transmitter

To illustrate the principle of MTCM, an example of 2-state MTCM was discussed. Figure given earlier is the 2- state trellis diagram for rate 1/2 conventional trellis coded QPSK and the corresponding multiple trellis diagram for the same coded modulation was given next. In MTCM case $m = 1$ and $k = 2$ and thus there are $2^{nk} = 4$ branches going from each state. With 2-state trellis this means that there must be two parallel branches between each pair of states. Also since $k = 2$ then there are two output QPSK symbols assigned to each branch.

Asymmetric Constellations –MTCM

Partition the original signal point constellation into two sub-constellations each with maximum distance among its signal points. Signals in partition 1 are assigned to transitions.



$$d_1^2 = 4\,sin^2\frac{f}{2} = 2(1 - cos\ f)$$

$$d_2^2 = 2$$

$$d_3^2 = 4\,sin^2\left(\frac{p}{4} + \frac{f}{2}\right) = 2(1 + sin\ f)$$

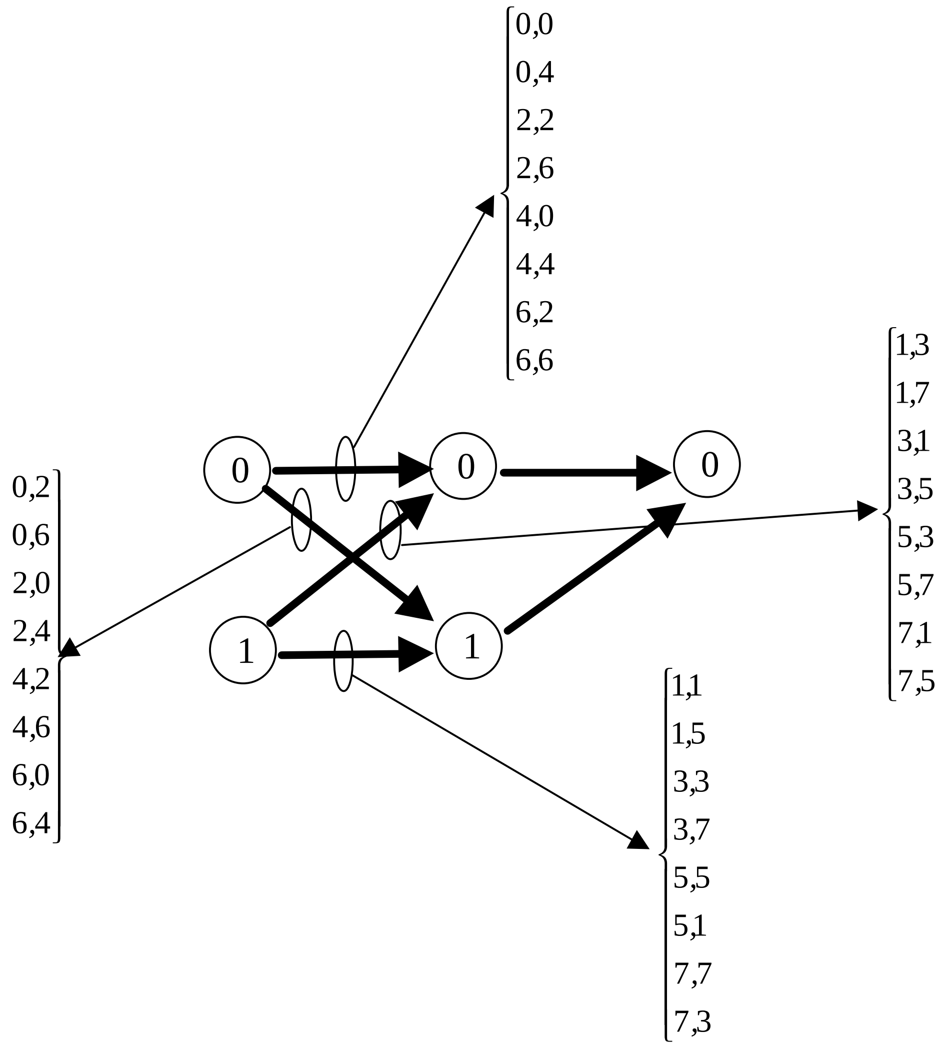
$$d_4^2 = 4$$

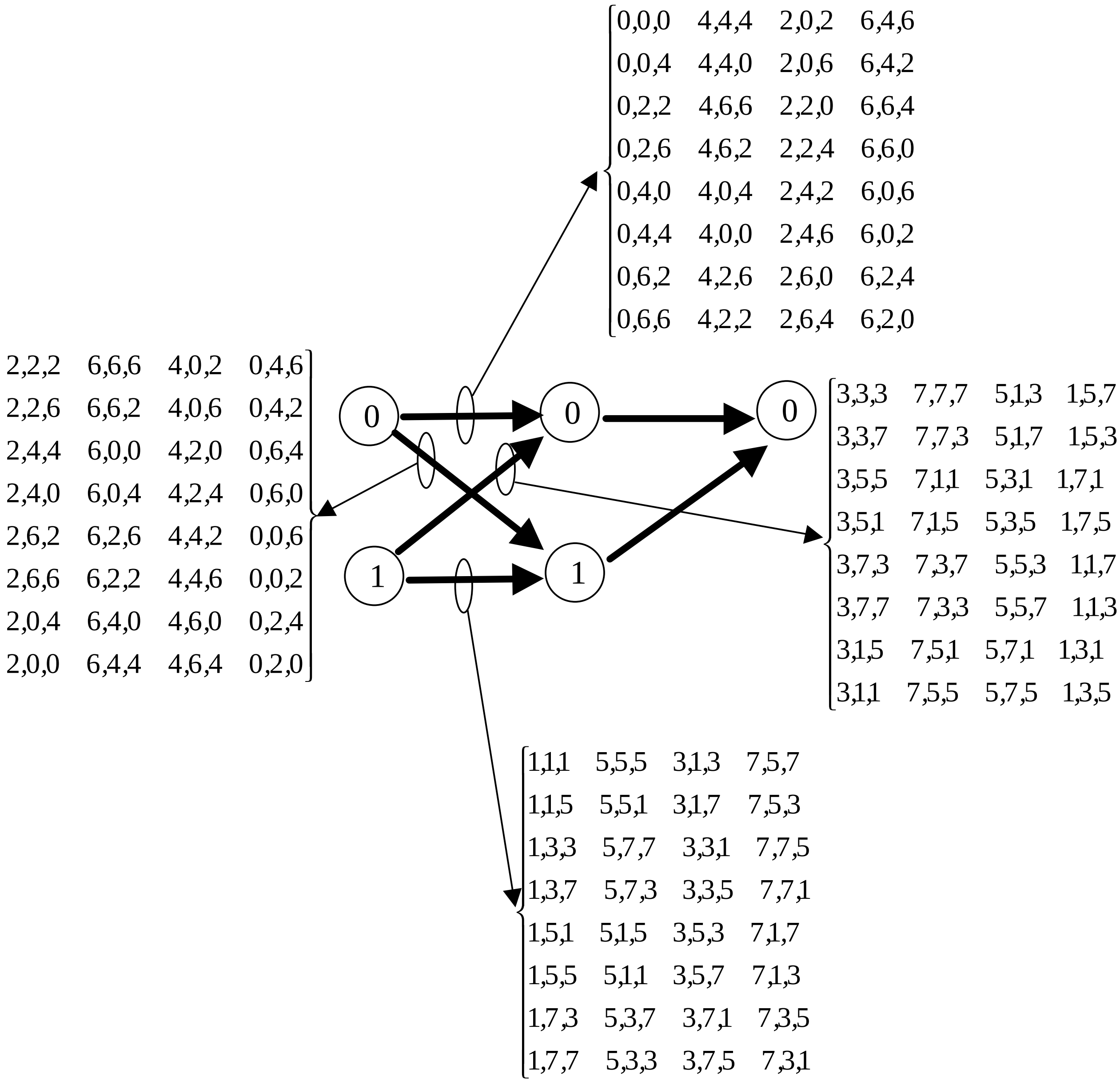
$$d_5^2 = 4\,sin^2\left(\frac{p}{2} - \frac{f}{2}\right) = 2(1 + cos\ f)$$

$$d_6^2 = 2$$

$$d_7^2 = 4\,sin^2\left(\frac{p}{4} - \frac{f}{2}\right) = 2(1 - sin\ f)$$

Two state trellises for multiplicity 2 and 3 are shown below.


0,0
0,4
2,2
2,6
4,0
4,4
6,2
6,6
1,3
1,7
3,1
3,5
5,3
5,7
7,1
7,5
0,2
0,6
2,0
2,4
4,2
4,6
6,0
6,4
0
0
0
1
1
1,1
1,5
3,3
3,7
5,5
5,1
7,7
7,3


0,0,0 4,4,4 2,0,2 6,4,6
0,0,4 4,4,0 2,0,6 6,4,2
0,2,2 4,6,6 2,2,0 6,6,4
0,2,6 4,6,2 2,2,4 6,6,0
0,4,0 4,0,4 2,4,2 6,0,6
0,4,4 4,0,0 2,4,6 6,0,2
0,6,2 4,2,6 2,6,0 6,2,4
0,6,6 4,2,2 2,6,4 6,2,0
2,2,2 6,6,6 4,0,2 0,4,6
2,2,6 6,6,2 4,0,6 0,4,2
2,4,4 6,0,0 4,2,0 0,6,4
2,4,0 6,0,4 4,2,4 0,6,0
2,6,2 6,2,6 4,4,2 0,0,6
2,6,6 6,2,2 4,4,6 0,0,2
2,0,4 6,4,0 4,6,0 0,2,4
2,0,0 6,4,4 4,6,4 0,2,0
0
0
0
1
1
3,3,3 7,7,7 5,1,3 1,5,7
3,3,7 7,7,3 5,1,7 1,5,3
3,5,5 7,1,1 5,3,1 1,7,1
3,5,1 7,1,5 5,3,5 1,7,5
3,7,3 7,3,7 5,5,3 1,1,7
3,7,7 7,3,3 5,5,7 1,1,3
3,1,5 7,5,1 5,7,1 1,3,1
3,1,1 7,5,5 5,7,5 1,3,5
1,1,1 5,5,5 3,1,3 7,5,7
1,1,5 5,5,1 3,1,7 7,5,3
1,3,3 5,7,7 3,3,1 7,7,5
1,3,7 5,7,3 3,3,5 7,7,1
1,5,1 5,1,5 3,5,3 7,1,7
1,5,5 5,1,1 3,5,7 7,1,3
1,7,3 5,3,7 3,7,1 7,3,5
1,7,7 5,3,3 3,7,5 7,3,1