

ГИПЕРФАЗОВАЯ МОДУЛЯЦИЯ – ОПТИМАЛЬНЫЙ МЕТОД ПЕРЕДАЧИ ЦИФРОВЫХ СООБЩЕНИЙ

Быховский М.А., д.т.н., профессор Московского технического университета связи и информатики, e-mail: bykhmark@gmail.com.

HYPERPHASE MODULATION – THE OPTIMAL METHOD OF MESSAGE TRANSMISSION

Bykhovskiy M.

Author proposes a new method of message transmission by using Hyperphase Modulation. In HPM, all signal points (SP) of the corresponding signal ensemble (SE) are located on the surface of N-dimensional sphere and the distance (d) between the closest SPs with the increase in N dimension is measured according to the formula $d \sim \sqrt{N}$. Such SE, as was shown by Shannon in 1959, is optimal given large enough value of N and allows to ensure the transmission of signals in the telecommunication channel with a limited frequency band with a set speed of transmission and maximum possible noise immunity.

The article explores dependence of the speed of message transmission with the use of HPM on minimal phase distance between SPs in this multidimensional SE. It is shown that by applying HPM method it is possible to achieve a substantial net energy gain as compared with other two-dimensional SEs, such as QAM. Such gain grows with the increase of N and can be quite significant.

The author develops an effective, from a calculation standpoint, coding algorithm of the transmitted message – by using the transmitted message number (m) to calculate phase coordinates of signals' SPs. The values of those coordinates are then used for formation of transmitted signals with HPM. The major difference of this coding method from those that are used in modern telecommunication systems is that that method, while improving noise immunity, it does not contribute additional (excessive) symbols into a sequence of informational symbols of a transmitted message.

Key words: A new Method for Digital Modulation- Hyperphase Modulation (HPM), the Design of the Ensemble of Signals, Energy and Spectral Efficiency of HPM.

Ключевые слова: гиперфазовая модуляция, конструкция ансамбля сигналов, энергетическая и спектральная эффективность, многомерные поверхностно-сферические коды.

Введение

В фундаментальных работах К. Шеннона [1-2] для пояснения основных функций передачи и приема цифровых сообщений широко использовались геометрические представления. Было показано, что для передачи сообщений по каналу связи необходимо создать оптимальный ансамбль сигналов – АС (кодов), длительность в полосе частот F которых T . Каждому из сигналов, входящих в АС, может быть поставлена в соответствие сигнальная точка (СТ) в N – мерном пространстве ($N = 2FT$, FT – целое число).

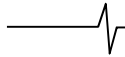
Было также установлено, что координаты этих точек должны быть выбраны так, чтобы расстояние между ближайшими СТ было не меньше заданного значения d .

Представление сигналов в виде СТ в N -мерном сигнальном пространстве основано на теореме Найквиста-Котельникова-Шеннона, согласно которой в канале связи с полосой частот F , любые сигналы могут быть представлены в виде суммы $N = 2FT$, элементарных составляющих. Безразмерную полосу частот канала связи с заданным параметром FT можно считать нормированной длительностью передаваемых сигналов, для геометрического представления которых удобно использовать их разложение по базису ортогональных функций.

Настоящий материал посвящен изложению проблем построения конкретного оптимального ансамбля сигналов (кодов) в многомерном пространстве состоит из трех частей. В данной статье описана процедура построения многомерных поверхностно-сферических кодов, оптимальность которых была доказана Шенноном. При формировании сигналов, принадлежащих этому коду, осуществляется фазовая модуляция $N-1$ генераторов элементарных гармонических колебаний, входящих в состав модулятора. Поэтому метод передачи сообщений с помощью поверхностно-сферических кодов логично называть гиперфазовой модуляцией (ГПФМ, НРМ – Hyperphase Modulation). Разработан алгоритм, ставящий в соответствие номеру m передаваемого сообщения (алгоритм кодирования сообщений) сигналы $S_m(t)$, входящие в АС. Вычислительная сложность этого алгоритма растет лишь линейно с увеличением N .

На практике наиболее часто для этого применяются гармонические функции и при этом, как правило, используют методы временного и частотного уплотнения каналов.

При передаче сообщений по каналу связи с полосой F методом временного уплотнения (TDMA – Time Division Multiple Access) интервал T , делится на подинтервалы, длительность которых равна $\tau_T = T / (FT)$ и при $t \in T_i = [(i-1)\tau_T, i\tau_T]$ ($i = 1...FT$) в канале связи передаются элементарные сигналы $V_{(2i-1),m}(t) = x_{m,(2i-1)} \times \sin \omega(t - T_i)$ и $V_{2i,m}(t) = x_{m,2i} \cos \omega(t - T_i)$, за пределами



интервала T равные 0. Параметры $x_{m,(2i-1)}$ и $x_{m,2i}$ определяются передаваемой цифровой информационной последовательностью символов на входе модулятора. Сигнал на входе приемника за интервал T может быть

$$S_m(t) = \sqrt{2NP_S} \left[\sum_{i=1}^{FT} V_{(2i-1),m}(t) + V_{2i,m}(t) \right],$$

где P_S – средняя мощность одного элементарного сигнала на передающем конце линии связи.

Другим методом передачи сигналов является частотное уплотнение канала связи ортогональными поднесущими (OFDM – Orthogonal Frequency-Division Multiplexing). Длительность передачи каждой поднесущей равна полной длительности сигнала T , на каждой из них передаются элементарные сигналы $V_{(2i-1),m}(t) =$

$$= x_{m,(2i-1)} \sin \omega_i t \text{ и } V_{2i,m}(t) = x_{m,2i} \cos \omega_i t, \text{ частоты которых выбраны так, чтобы выполнялись условия } \int_0^T V_{i,m}(t) V_{j,m}(t) dt = 0 \text{ при } i \neq j. \text{ Для этого необходимо,}$$

чтобы $|\omega_i - \omega_j| = 2\pi|i - j|/T$ при $i \neq j$. В данном случае сигнал, принятый за время T , также может быть представлен выше приведенным соотношением.

Отметим, что оба метода (TDMA и OFDM) с точки зрения геометрической трактовки алгоритмов передачи и приема сигналов полностью эквивалентны: передаче сигналов во временном интервале T_i в системе TDMA соответствует передача таких же сигналов на частоте ω_i . Однако следует отметить, что физические параметры этих сигналов существенно отличаются. Кроме того, как было показано С. Вейнштейном и П. Эбертом [3], используя быстрое преобразование Фурье, эти сигналы могут быть преобразованы из одного вида в другой.

Шенноном было показано, что CT оптимального ансамбля сигналов $S_m(t) (m = 1 \dots M)$ при больших значениях N должны быть расположены на поверхности N -мерной сферы. Отметим, что в частном случае, когда $N = 2$, двумерная сфера представляет собой окружность и в таком случае принято, что для передачи сигналов используется AC с многофазной модуляцией. При любых значениях N для всех входящих в AC поверхность-сферических сигналов $S_m(t)$, имеющих длительность T , справедливо соотношение

$$\int_0^T S_m^2(t) \frac{dt}{T} = NP_S, \tag{1}$$

где P_S – мощность сигнала, приходящаяся на одно измерение N -мерного пространства, в котором размещены CT , соответствующие AC .

Параметры $x_{m,(2i-1)}$ и $x_{m,2i}$ могут рассматриваться как нормированные евклидовы координаты CT , соответствующие сигналу $S_m(t)$. Из (1) следует, что нормированные координаты всех CT удовлетворяют условию

$$\sum_{i=1}^{FT} [x_{m,(2i-1)}^2 + x_{m,2i}^2] = 1, \tag{2}$$

т.е. при нормировке все CT лежат на поверхности N -мерной сферы с радиусом $R_S = 1$.

На рис. 1 показана структура систем цифровой системы связи [1, 2], фундаментальные исследования которой впервые были начаты Шенноном. Эта система на передающей стороне имеет последовательно соединенные кодер, на вход которого поступает $n \approx \log_2 M$ двоичных информационных символов, где M – количество передаваемых сигналов.

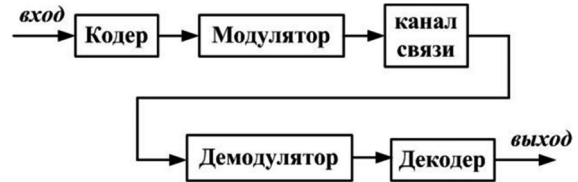


Рис. 1. Схема системы связи

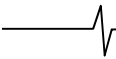
Передаваемые блоки информационных символов по сути представляют собой номер m CT , соответствующий сигналу $S_m(t)$, принадлежащему используемому для передачи сообщений AC . По Шеннону в кодере осуществляется формирование N индексов фазовой модуляции, определяющих значения координат CT , соответствующих блоку входных информационных символов. В модуляторе формируется аналоговый сигнал, принадлежащий AC , выбранному для передачи сообщений. Оптимальный AC должен быть сконструирован таким образом, чтобы евклидовы расстояния между двумя ближайшими CT , соответствующими входящими сигналам, были бы равны d . От значения d зависит количество составляющих в AC и, следовательно, скорость передачи сообщений.

На приемной стороне линии связи принятый сигнал поступает на вход демодулятора, основной функцией которого является определение координат CT , соответствующей индексам модуляции принятого сигнала. Значения этих координат (индексов модуляции) подаются на вход декодера. В оптимальном демодуляторе по Шеннону осуществляется определение координат CT , принадлежащей одному из передаваемых сигналов в N -мерном евклидовом сигнальном пространстве, наиболее близко расположенному к CT , соответствующий принятому сигналу, искаженному шумом, действующим на входе демодулятора.

Функциональным назначением декодера является формирование блока двоичных информационных символов, определяющих номер m принятого сообщения, используя для этого определенные в демодуляторе координаты CT .

Следует отметить, что, согласно Шеннону [1, 2] в оптимальной системе передачи сообщений по существу осуществляются только модуляция передаваемых сигналов (формирование сигналов $S_m(t)$), а помехоустойчивое кодирование, при котором по каналу связи передаются не только информационные, но и избыточные символы, не используется.

Шенноном было установлено, что, применяя оптимальный AC , можно достичь предельно высокой помехоустойчивости приема сигналов в том случае, когда



скорость передачи сообщений $R = \log_2 M / T$ меньше пропускной способности канала связи:

$$C = F \cdot \log_2(1 + \rho_0) \quad (3)$$

где $\rho_0 = P_{S0} / P_n$ – отношение средней мощности принятого сигнала (P_{S0}) к мощности шума (P_n), действующего на входе демодулятора.

При проектировании системы связи полоса частот канала, по которому передаются сообщения, обычно является заданной. Введем важные нормированные параметры, к значениям которых предъявляются требования при проектировании системы связи. Создаваемая система должна обеспечивать заданную удельную скорость передачи сообщений ($R_f = (R / F) = \frac{\log_2 M}{N / 2}$), а канал связи должен иметь удельную пропускную способность ($C_f = C / F$). Отметим, что между параметрами ρ_0 и C_f имеется соотношение $\rho_0 = (2^{C_f} - 1)$ и что условие $R < C$ эквивалентно условию $R_f < C_f$. К системе связи предъявляется также требования обеспечения необходимой надежности связи – значения вероятности ошибочного приема сигналов ($P_{er}(\rho_0, FT)$).

Как было показано Шенноном [1-2], при выполнении условия $R_f < C_f$ вероятность ошибочного приема сигналов ($P_{er}(\rho_0, FT)$) может быть сколь угодно малой при увеличении длительности сигналов, входящих в оптимальный АС. Применяя такой АС, можно при $FT \rightarrow \infty$ обеспечить предел $P_{er}(\rho_0, FT) \rightarrow 0$. В работе Шеннона [2] были получены зависимости $P_{er}(\rho_0, FT)$ от ρ_0 и FT , из которых при $P_{er}(\rho_0, FT) = const$ могут быть определены соотношения между ρ_0 и FT , позволяющие выбрать размерность N -мерного сигнального пространства, в котором размещены CT сигналов оптимального АС.

В [4, 5] разработаны простые методы определения зависимостей $P_{er}(\rho_0, FT)$, относящиеся к классам поверхностно- и объемно-сферических АС (в литературе используются также эквивалентный термин – сигнално-кодовые конструкции), которые приводят к результатам Шеннона [2]. Из [2, 4, 5] следует, что в оптимальном АС параметр d , определяющий расстояние между ближайшими CT в АС, с увеличением N изменяется пропорционально \sqrt{N} ($d \sim \sqrt{N}$).

В работах Шеннона, посвященных помехоустойчивым кодам, было показано, что в N -мерном пространстве возможно построить ансамбль сигналов, в котором евклидово расстояние между любыми двумя сигналами имеет максимально возможное значение, зависящее только от значений FT , ρ_0 и R_f .

Термин «кодирование» по Шеннону существенно отличается от современного понятия в теории кодирования. К поверхностно- и объемно-сферическим ансамблям сигналов, исследованным Шенноном, широко используемое в этой теории понятие «избыточность», во-

обще не применимо.

Отметим, что в современных системах связи для передачи кодовых комбинаций, сформированных в кодере, широкое применение находят АС с квадратурной амплитудной модуляцией (QAM – Quadrature Amplitude Modulation) и с амплитудно-фазовой модуляцией (АФМ – Amplitude and Phase Modulation). Находят применение также и другие виды АС, в частности сигналов с квадратурно-фазовой (QPSK – Quadrature phase-shift keying), с частотной модуляцией (FSK – Frequency-shift keying) и др. [6, 7]. Все указанные АС имеют достаточно простую структуру и относятся к классу двумерных. Принадлежащие им сигналы представляют собой простое гармоническое колебание, модулированное по амплитуде, фазе или частоте. Следует указать, что в классе двумерных АС обеспечить между любыми сигналами достаточно большое значение минимального расстояния d , которое необходимо для их демодуляции с малой вероятностью ошибки, можно только за счет увеличения мощности P_s передаваемого сигнала. Поэтому в современных системах связи получение малой вероятности ошибки при демодуляции сигналов является проблематичным.

Во всех двумерных АС при заданном значении P_s величина d уменьшается с увеличением R_f . Поэтому, используя двумерные АС, невозможно получить высокую помехоустойчивость и скорость передачи сообщений, которую может обеспечить идеальный по Шеннону АС при $d \sim \sqrt{N}$.

Во многих современных системах передачи сообщений кодирование и модуляция осуществляются раздельно. При этом кодер (рис. 1), построенный на основе современной теории кодирования, осуществляет формирование кодовых комбинаций, имеющих длину n и содержащих как информационные символы, поступившие на его вход, так и дополнительные избыточные символы, связанные с информационными определенными математическими соотношениями. Включение в каждую кодовую комбинацию избыточных символов увеличивает расстояние между разными кодовыми комбинациями и позволяет исправлять в них ошибки, возникающие из-за действия шумов при приеме сигналов. Однако избыточные символы уменьшают реальную скорость передачи сообщений. Символы, сформированные в кодере, поступают на вход модулятора. В соответствии с блоком двоичных символов длины $k < n$ ($k \in 1, \dots, 10$) в модуляторе формируется один из 2^k аналоговых сигналов, принадлежащих АС для передачи сообщений в канал связи.

Важным вкладом в теорию связи стали работы Унгербоэка [8], разработавшего методы так называемой решетчатой модуляции (Treillis-Coded Modulation – TCM), предназначенной для передачи сообщений в каналах связи, ограниченных по полосе частот. Этот вид модуляции предусматривает введение избыточности в передаваемую информационную последовательность символов для формирования N -мерных АС, элементами которых являются двумерные M -позиционные АС типа QAM или PAM. Избыточные символы используются та-

ким образом, чтобы из M^N возможных CT , соответствующих всем комбинациям координат CT для N последовательно передаваемых в канал связи элементарных двумерных сигналов, обеспечивалась бы передача только тех, для которых расстояние между ближайшими CT превышает заданное расстояние между ближайшими CT в двумерных AC значение d . Энергетический выигрыш от использования TCM вместо QAM без применения помехоустойчивого кодирования при $M=128$ и 256 достигает примерно 6 дБ [6]. При этом скорость передачи сообщений уменьшается до $2/3$ от той, которая могла быть обеспечена при передаче этих же сообщений с помощью QAM без введения избыточных символов. Исследованию для передачи сообщений TCM и других видов многомерных AC посвящен также ряд обстоятельных исследований [9-11].

В ряде работ были выполнены многочисленные исследования, связанные с выяснением возможности применения для передачи сообщений многомерных сигналов. В этих исследованиях оценивались зависимости скорости передачи сообщений с помощью многомерных AC от их длительности и от значения минимального расстояния между ними, были предложены алгоритмы построения AC для сравнительно небольших значений N . Результаты таких исследований приведены, например, в [12], где даны обширные ссылки на соответствующую литературу.

Следует отметить, что общих решений задачи построения многомерного AC, с помощью которого сообщения передаются с потенциально возможной [2] помехоустойчивостью, в них не предлагалось.

Как видно из рис. 1, на приемной стороне сигналы поступают на вход демодулятора, в котором при традиционном построении системы связи в результате демодуляции сигнала формируются блоки из n символов кодовой комбинации, соответствующей принятому сигналу. Часть этих символов может быть принята с ошибкой. Эти кодовые комбинации поступают на вход декодера, в котором осуществляется исправление ошибок, появившихся на выходе демодулятора, и формируется последовательность информационных символов, являющихся копией переданной информации. Для обеспечения высокой помехоустойчивости приема – малой вероятности ошибки в информационной последовательности, применяются достаточно длинные коды ($n \gg 1$). При их декодировании используются весьма сложные алгоритмы. Современная теория кодирования предлагает для применения в телекоммуникационных системах ряд кодов и методов их декодирования. Наиболее развиты и широко применяются линейные коды, алгоритмы кодирования и декодирования которых основаны на теории Галуа, а также коды Галлагера, сверточные и турбо-коды [6, 7, 13].

Следует отметить, что в работах Шеннона не рассматривались вопросы построения конкретных AC в многомерном пространстве. Построение таких AC требует разработки алгоритма определения координат CT , соответствующих сигналам оптимального AC, таким образом, чтобы при заданном значении R_f наименьшее расстояние d между сигналами, входящими в AC, уве-

личивалась бы с увеличением N по закону $d \sim \sqrt{N}$ – также, как идеально по Шеннону AC.

В работах Шеннона не затрагивались также вопросы, связанные с разработкой алгоритма нумерации сигналов, входящих в AC. Однако важно обратить внимание на то, что процедуру кодирования для оптимального кода (AC), содержащего $M = 2^{R_f(FT)}$ сигналов, Шеннон трактовал, как операцию присвоения определенного номера m каждому из M возможных сигналов, считая, что двоичная запись этого номера представляет собой информационную последовательность знаков, которые должны быть переданы по каналу связи.

Гиперфазовый модулятор и кодер сообщений

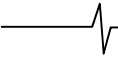
Процедуры построения гиперфазового модулятора, в котором определяются координаты всех CT ансамбля сигналов, размещенных на поверхности N -мерной сферы радиуса 1 и определения количества CT в AC можно представить в ниже следующем виде. Известно [14], что нормированные евклидовы координаты любой CT , соответствующей сигналу $S_m(t)$, расположенной на поверхности N -мерной сферы радиуса 1, могут быть выражены в через ее гиперсферические (фазовые) координаты с помощью следующих N соотношений:

$$\begin{aligned} x_{m1} &= \sin \varphi_{m1}, \quad x_{m2} = \cos \varphi_{m1} \sin \varphi_{m2}, \\ x_{mk} &= \left(\prod_{i=1}^{k-1} \cos \varphi_{mi} \right) \sin \varphi_{mk}, \dots \\ \dots x_{m(N-1)} &= \left(\prod_{i=1}^{N-2} \cos \varphi_{mi} \right) \sin \varphi_{m(N-1)}, \dots x_{mN} = \\ &= \left(\prod_{i=1}^{N-2} \cos \varphi_{mi} \right) \cos \varphi_{m(N-1)}. \end{aligned} \quad (4)$$

Для упрощения записи этих формул обозначим:

$$\begin{aligned} R_{m1} &= 1, \quad R_{m2} = \cos(\varphi_{m1}) \quad \text{и} \quad R_{mk} = \left(\prod_{i=1}^{k-1} \cos \varphi_{mi} \right). \quad \text{Тогда при-} \\ \text{веденные формулы примут более компактный вид:} \\ x_{m1} &= R_{m1} \sin \varphi_{m1}, \quad x_{m2} = R_{m2} \sin \varphi_{m2}, \\ x_{mk} &= R_{mk} \sin \varphi_{mk}, \dots x_{m(N-1)} = R_{m(N-2)} \sin \varphi_{m(N-1)}, \\ x_{mN} &= R_{m(N-2)} \cos \varphi_{m(N-1)}. \end{aligned} \quad (5)$$

Эти соотношения показывают, что при заданном значении P_s каждая CT в N -мерном пространстве определяется N евклидовыми координатами x_{mk} ($k = 1 \dots N$), зависящими от $(N-1)$ параметров φ_{km} – ее фазовых координат. Эти параметры могут быть использованы для передачи сообщений в качестве информационных. Особенностью представления координат CT , расположенных на поверхности гиперсферы, является их взаимозависимость: для сигнала $S_m(t)$ евклидова координата по оси x_1 зависит только от одной фазовой координаты φ_{m1} , координата по оси x_2 – от двух фазовых координат φ_{m1} и φ_{m2} , координата по k -ой ортогональной оси N -мерного пространства – от фазовых координат $\varphi_{m1}, \varphi_{m2}, \dots, \varphi_{m(k-1)}$. Важно указать диапазон возможных значений параметров φ_{km} [14]:



$$\begin{aligned} & -0,5\pi \leq \varphi_{mk} \leq 0,5\pi, \text{ если } 1 \leq k \leq (N-2), \\ & \text{и } 0 \leq \varphi_{m(N-1)} \leq 2\pi. \end{aligned} \quad (6)$$

Из (5) следует, что если значения координат x_{mk} СТ для $1 \leq k \leq (N-2)$ заданы, то ее координаты $x_{m(N-1)}$ и x_{mN} лежат на окружности, радиус которой зависит от значений x_{mk} ($k=1, \dots, (N-2)$) и равен $R_{m(N-2)}$.

Алгоритм синтеза АС, предназначенного для передачи сообщений по каналу связи с удельной скоростью R_f и высокой помехоустойчивостью, состоит в выборе фазовых координат φ_{mk} таким образом, чтобы расстояние между ближайшими СТ, размещенными на поверхности N -мерной сферы, было бы равно заданному значению d . Кроме того, количество сигналов в N -мерном АС (и соответствующих им СТ) должно быть таким, чтобы этот ансамбль позволял передавать сообщения по каналу связи с заданной удельной скоростью R_f . Для этого количество СТ, размещенных на поверхности N -мерной сферы, должно быть равно $M = 2^{R_f(N/2)}$.

При синтезе АС необходимо, чтобы каждый из информационных параметров (фазовых координат СТ) мог изменяться дискретно, т.е. $\varphi_{mk} = I_{mk} \Delta \varphi_{mk}$, где I_{mk} – целые числа ($-M_k \leq I_{mk} \leq M_k$) – индексы модуляции передаваемого сообщения, соответствующие его номеру, однозначно определяющие значение m . Операции вычисления индексов выполняются на передающем конце линии связи в блоке их формирования, который можно назвать кодером. Формирование сигнала $S_m(t)$, зависящего от фазовых координат φ_{mk} ($k=1 \dots (N-1)$), осуществляется в гиперфазовом модуляторе.

Процедура формирования сигналов ГПФМ определяется формулами (4) и (5).

Ниже показано, что значения евклидовых координат СТ для АС, их количество (M), а также значения $\Delta \varphi_{mk}$ полностью определяются выбором только одного параметра $\Delta \varphi_1$, причем значения $\Delta \varphi_{mk}$ и M_{mk} зависят не только от $\Delta \varphi_1$, но и, как следует из (4) и (5), от того, какие значения присвоены фазовым координатам φ_{mi} при $i \leq (k-1)$.

Найдем соотношения, определяющие $\Delta \varphi_{mk}$ и M_{mk} , используя для этого условие минимального евклидова расстояния d между любыми двумя соседними СТ, соответствующими принадлежащим АС сигналам $S_{m_1}(t)$ и $S_{m_2}(t)$:

$$\begin{aligned} d^2 &= \int_0^T [S_{m_1}(t) - S_{m_2}(t)]^2 \frac{dt}{T} = \\ &= NP_S \left\{ \sum_{i=1}^{FT} [x_{m_1, (2i-1)} - x_{m_2, (2i-1)}]^2 + [x_{m_1, 2i} - x_{m_2, 2i}]^2 \right\}. \end{aligned} \quad (7)$$

Нормируя расстояние d к $\sqrt{NP_S}$ ($d_0 = d / \sqrt{NP_S}$), соотношение (7) можно записать:

$$d_0^2 = \left[\sum_{i=1}^{FT} (x_{m_1, (2i-1)} - x_{m_2, (2i-1)})^2 + (x_{m_1, 2i} - x_{m_2, 2i})^2 \right]. \quad (8)$$

Значения k -х фазовых координат (k -х информационных параметров) в рассматриваемом АС могут принимать значения, равные $\varphi_{mk} = I_{mk} \Delta \varphi_k$, где I_{mk} – целые положительные или отрицательными числа. Возможный диапазон их изменений будет определен ниже.

Отметим, что минимальное фазовое расстояние между двумя сигналами $S_{m_1}(t)$ и $S_{m_2}(t)$, отличающимися друг от друга только k -й угловой координатой, равно $\Delta \varphi_{mk}$ (при этом $\min |I_{m_1 k} - I_{m_2 k}| = 1$). Выполнив в (6) выкладки с учетом (5), найдем нормированное евклидово расстояние между ними:

$$d_0^2 = 4 \sin^2(\Delta \varphi_1 / 2) = 4 \left(\prod_{i=1}^{k-1} \cos \varphi_{mi} \right)^2 \sin^2(\Delta \varphi_{mk} / 2) \quad (9)$$

Второе соотношение в (9) относится к случаю, когда у сигналов $S_{m_1}(t)$ и $S_{m_2}(t)$ значения первых фазовых координат φ_{mi} совпадают при $i \leq (k-1)$, и отличаются только значения $\varphi_{m_1 k}$ и $\varphi_{m_2 k}$.

Используя уравнения (9), получим все основные параметры АС: значения φ_{mk} всех фазовых координат СТ, значения M_{mk} , определяющие возможный диапазон изменений параметра I_{mk} для каждого значения k , количество СТ для рассматриваемого АС (M), а также удельную скорость передачи сообщений R_f , которая может быть обеспечена этим АС.

Из (9) следуют соотношения, определяющие зависимость $\Delta \varphi_{mk}$ и M_{mk} от $\Delta \varphi_1$:

$$\left(\prod_{i=1}^{k-1} \cos \varphi_{mi} \right)^2 \sin^2(\Delta \varphi_{mk} / 2) = \sin^2(\Delta \varphi_1 / 2). \quad (10)$$

Рассмотрим построение АС для случая, когда $N \geq 3$. Формула (10) позволяет последовательно находить значения $\Delta \varphi_{mk}$ в зависимости от $\Delta \varphi_1$, а также значения M_{mk} , определяющие возможный диапазон изменения параметров I_{mk} . Величины $\Delta \varphi_{mk}$ равны:

$$\begin{aligned} \Delta \varphi_{mk}(\Delta \varphi_1, \vec{I}_{m(k-1)}) &= \\ &= 2 \arcsin \left[\left| \sin(\Delta \varphi_1 / 2) \right| / \prod_{i=1}^{k-1} \left| \cos(I_{mi} \Delta \varphi_{mi}) \right| \right], \end{aligned} \quad (11)$$

Отметим, что значения $\Delta \varphi_{mk}$ зависят только от абсолютного значения индекса $|I_{m1}|$. Запись формул более компактна, если указать на зависимость индексов I_{mi} от многих параметров ($i=1 \dots k$), в (11) и в последующих формулах строка из k символов I_{mk} записывается в виде $\vec{I}_{mk} = (I_{m1}, I_{m2}, \dots, I_{mk})$.

Определить из (11) значение $\Delta \varphi_{mk}(\Delta \varphi_1, \vec{I}_{m(k-1)})$ можно только в том случае, если выполняется условие

$$|\sin(\Delta\varphi_1 / 2)| / \prod_{i=1}^{k-1} |\cos(I_{mi} \Delta\varphi_{mi})| \leq 1. \quad (12)$$

Следовательно, целые числа I_{mk} при любом значении k должны лежать в интервале $-M_k(\Delta\varphi_1, \bar{I}_{m(k-1)}) \leq I_{mk} \leq M_k(\Delta\varphi_1, \bar{I}_{m(k-1)})$, где $M_1(\Delta\varphi_1, \bar{I}_{m(k-1)}) = \text{int}\{\arccos[|\sin(\Delta\varphi_1 / 2)| / \Delta\varphi_1]\}$ и при $k \geq 2$

$$M_k(\Delta\varphi_1, \bar{I}_{m(k-1)}) = \text{int} \left\{ \frac{\arccos \left[|\sin(\Delta\varphi_1 / 2)| / \left(\prod_{i=1}^{k-1} |\cos(I_{mi} \Delta\varphi_{mi})| \right) \right]}{\Delta\varphi_{mk}} \right\}. \quad (13)$$

Формулы (11)-(13), где $\text{int}(x)$ – целая часть числа x , определяют все фазовые координаты CT , т.е. их евклидовы координаты. Знание $\Delta\varphi_{mk}(\Delta\varphi_1, \bar{I}_{m(k-1)})$ и $M_k(\Delta\varphi_1, \bar{I}_{m(k-1)})$ позволяет составить схему гиперфазового модулятора, указанного на функциональной схеме системы связи на рис. 1.

Формулы (11) и (13) показывают, что каждая фазовая координата может принимать разные значения в зависимости от чисел, входящих в состав строки $\bar{I}_{m(k-1)} = (I_{m1}, I_{m2} \dots I_{m(k-1)})$.

Важно отметить, что целые числа I_{mk} в интервале $-M_k(\Delta\varphi_1, \bar{I}_{m(k-1)}) \leq I_{mk} \leq M_k(\Delta\varphi_1, \bar{I}_{m(k-1)})$, могут рассматриваться как индексы вложенных подмножеств разного порядка, на которые может быть разбито множество всех CT , размещенных на поверхности N -мерной сферы. Подмножества сигнальных точек 1-го порядка, которым присваиваются индексы $|I_{m1}|$ (для краткости обозначены далее $\text{Пм}_{|I_{m1}|}$), отличаются друг от друга тем, что принадлежащие им CT имеют разные абсолютные значения 1-й фазовой координаты $\varphi_1(I_{m1}) = I_{m1} \Delta\varphi_1$, зависящей от индекса I_{m1} . Количество таких подмножеств $\text{Пм}_{|I_{m1}|}$ равно $G_1 = 2M_1(\Delta\varphi_1) + 1$, где $M_1(\Delta\varphi_1) = \text{int}\{\arccos[|\sin(\Delta\varphi_1 / 2)| / \Delta\varphi_1]\}$ – число значений, которые может принимать индекс $|I_{m1}|$. Каждое из подмножеств 1-го порядка с индексом $|I_{m1}|$ может быть разбито на подмножества 2-го порядка, которые обозначены $\text{Пм}_{|I_{m1}|, |I_{m2}|}$ и отличаются друг от друга тем, что принадлежащие им CT имеют разные значения 2-й фазовой координаты $\varphi_2(\bar{I}_{m2}) = I_{m2} \Delta\varphi_2(\Delta\varphi_1, |I_{m1}|)$ при $|I_{m1}| = \text{const}$. Количество таких подмножеств равно $G_2(I_{m1}) = 2M_2(\Delta\varphi_1, |I_{m1}|) + 1$. В общем случае, когда определены фазовые координаты CT $\varphi_1(I_{m1}), \varphi_2(\bar{I}_{m2}), \dots, \varphi_{mk}(\bar{I}_{mk}) = I_{mk} \Delta\varphi_{mk}(\Delta\varphi_1, \bar{I}_{m(k-1)})$, входящих в одно из подмножеств k -го порядка – $\text{Пм}_{|I_{m1}|, \dots, |I_{mk}|}$, то это подмножество может быть разбито на $G_{(k+1)}(\bar{I}_{mk}) = 2M_{(k+1)}(\Delta\varphi_1, \bar{I}_{mk}) + 1$ подмножеств $(k+1)$ -го порядка.

В каждое из этих подмножеств входят CT , у которых $(k+1)$ -я фазовая координата определяется формулой $\varphi_{m(k+1)}(\bar{I}_{m(k+1)}) = I_{m(k+1)} \Delta\varphi_{mk}(\Delta\varphi_1, \bar{I}_{mk})$. При этом $-M_{(k+1)}(\Delta\varphi_1, \bar{I}_{mk}) \leq I_{m(k+1)} \leq M_{(k+1)}(\Delta\varphi_1, \bar{I}_{mk})$ и $|I_{mi}| = \text{const}$ при $i \leq k$.

Отметим важные соотношения между значениями $\Delta\varphi_{mk}(\Delta\varphi_1, \bar{I}_{mk})$ и $M_k(\Delta\varphi_1, \bar{I}_{m(k-1)})$, которые следуют из (11) и (13):

$$\Delta\varphi_{mk}(\Delta\varphi_1, \bar{I}_{mk}) \geq \Delta\varphi_{m(k+1)}(\Delta\varphi_1, \bar{I}_{m(k+1)}) \quad (14)$$

$$\text{и } M_{(k+1)}(\Delta\varphi_1, \bar{I}_{mk}) \leq M_k(\Delta\varphi_1, \bar{I}_{m(k-1)}).$$

Из (14) следует, что с увеличением k увеличиваются значения $\Delta\varphi_{mk}(\Delta\varphi_1, \bar{I}_{mk})$, а количество подмножеств, на которое может быть разбито подмножество k -го порядка, уменьшается. Максимальное количество таких подмножеств $\text{Пм}_{|I_{m1}|}$ равно $M_1(\Delta\varphi_1) = \text{int} \left\{ \arccos \left[\frac{\sin \left(\frac{\Delta\varphi_1}{2} \right)}{\Delta\varphi_1} \right] \right\}$.

$$\text{В том случае, если } \Delta\varphi_1 \ll 1, \text{ то } M_1(\Delta\varphi_1) \approx \text{int}[\pi / 2\Delta\varphi_1].$$

Сигнальные точки, входящие в состав подмножеств $(N-1)$ -го порядка, у которых $(N-1)$ -я фазовая координата равна $\varphi_{m(N-1)} = I_{m(N-1)} \Delta\varphi_{m(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)})$, при заданных значениях $\bar{I}_{m(N-2)}$ лежат на двумерных окружностях. Так как для таких CT должно выполняться условие $0 \leq \varphi_{m(N-1)} \leq 2\pi$, то индексы $I_{m(N-1)}$ должны лежать в интервале

$$0 \leq |I_{m(N-1)}| \leq M_{(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)}) = \text{int}(2\pi / |\Delta\varphi_{m(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)})|), \quad (15)$$

где $M_{(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)})$ – количество CT , лежащих на одной окружности.

Формулы (13) и (15) позволяют определить общее количество CT для AC в N -мерном пространстве:

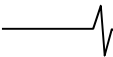
$$M_{sN}(\Delta\varphi_1) = \sum_{I_{m1}=-M_1(\Delta\varphi_1)}^{M_1(\Delta\varphi_1)} \dots \sum_{I_{m(N-1)}=-M_{(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)})}^{M_{(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)})} M_{(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)}), \quad (16)$$

а также удельную скорость передачи сообщений $R_f = [2\Delta \log_2 M_{sN}(\Delta\varphi_1)] / N. \quad (17)$

Основные характеристики сигналов с гиперфазовой модуляцией

С помощью формул (16) и (17) на рис. 2 построены зависимости удельной скорости передачи сообщений $R_f(\Delta\varphi_1)$ от параметра $\Delta\varphi_1$ (град) для исследованного AC , а также для AC , рассмотренного в работе Шеннона [2].

Для многомерных поверхностно-сферических AC зависимости $R_f(\Delta\varphi_1)$ построены при значениях $N = 3, 4$, и 6. Из рис. 2 видно, что кривые $R_f(\Delta\varphi_1)$ от значения N зависят довольно слабои при $N > 6$ они практически



совпадают с кривой при $N=6$. Зависимость $R_f(\Delta\varphi_1)$ от $\Delta\varphi_1$ гораздо более существенна. Для передачи сообщений с высокой удельной скоростью следует выбирать АС, у которых параметр $\Delta\varphi_1$ имеет небольшое значение.

Представляет интерес вопрос о выигрыше в увеличении минимального расстояния d между соседними СТ в N -мерном ансамбле сигналов с ГПФМ по отношению к ансамблю сигналов с QAM, который сегодня широко применяются на практике. Это связано с тем, что помехоустойчивость сигналов любого многопозиционного ансамбля зависит от минимального расстояния d между его СТ, при этом, чем больше это расстояние, тем выше помехоустойчивость приема сигналов.

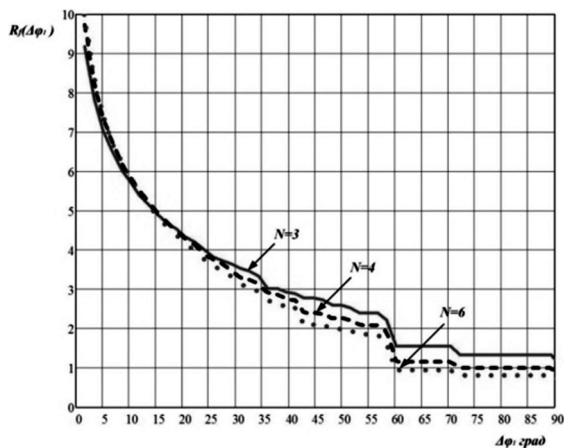


Рис. 2. Зависимости удельной скорости передачи сообщений $R_f(\Delta\varphi_1)$

Такой выигрыш определяется следующей формулой [6]

$$\mu(R_f, N) = 10 \log(d_{AC}^2 / d_{QAM}^2) \quad (18)$$

По формуле (18) на рис. 3 построены графики зависимости $\mu(R_f, FT)$ при $R_f = 2, 4, 6$ и 8.

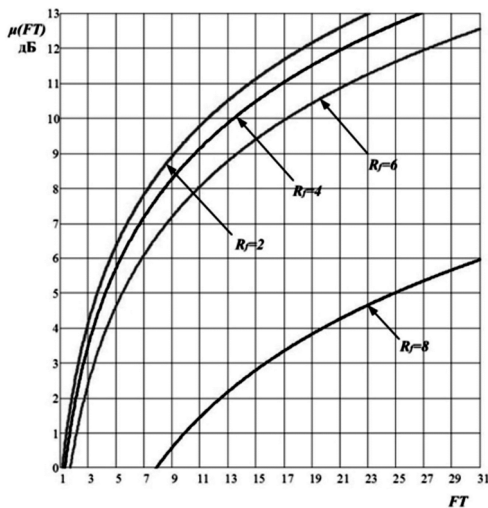


Рис. 3. Зависимости выигрыша $\mu(R_f, FT)$ от FT

В настоящее время в системах связи для передачи сообщений широко применяются, в частности, двумерные АС, такие, например, как QAM. Квадрат минимального расстояния между этой АС равен $d_{QAM}^2 = 6P_S \div \div(2^{R_f} - 1)$ [6], где R_f – удельная скорость передачи сообщений в системе связи, в которой применяется АС

с QAM. В исследованном N -мерном АС $d_{AC}^2 = = 4NP_S \{ \sin[\Delta\varphi_1(R_{fN}) / 2]^2 \}$. Квадрат минимального расстояния, как следует из (7), равен d_{AC}^2 и увеличивается с ростом N (в этой формуле зависимость $\Delta\varphi_1(R_f)$ является функцией, обратной $R_f(\Delta\varphi_1)$, представленной на рис. 2).

Из рис. 3 следует, что определенный (меньше 2 дБ) выигрыш в $\mu(R_f, N)$ от применения многомерного АС может быть получен даже при сравнительно небольших значениях $FT \cong 2$, если $R_f \leq 2 \div 6$. При больших значениях $R_f \geq 8$ от применения такого АС заметный выигрыш $\mu(R_f, N)$ может быть получен только в том случае, если $FT \geq 9$. Отметим, что, как известно [6], для двумерных АС сигналы с многофазной модуляцией (частный случай сигналов с ГПФМ) также уступают по своей энергетической эффективности сигналам QAM, если $R_f \geq 3$. Из рис. 3 также следует, что с увеличением R_f , для получения заметного выигрыша в $\mu(R_f, N)$ сигналов с ГПФМ по сравнению QAM, требуется увеличивать длительность применяемых многомерных сигналов.

Отметим, что при заданном $P_{er}(\rho_0, FT)$ – значении вероятности ошибочного приема сигналов, энергетический выигрыш системы связи с ГПФМ по сравнению с системами, использующими АС с QAM, меньше, чем $\mu(R_f, N)$, так как на прием сигналов, входящих в оптимальный АС ГПФМ, действуют шумы на каждую из N ортогональных составляющих этих сигналов.

Отметим, что Унгербоеком [6, 8] разработана решетчатая кодированная модуляция, сочетающая многопозиционную модуляцию и сверточное кодирование с избыточными символами. В ней при удельной скорости передачи сообщений $R_f \leq 4$ был получен выигрыш в $\mu(R_f, N)$ порядка 6 дБ [6, 8]. Для двумерных сигналов этот эффект определяет предельный энергетический выигрыш системы ТСМ по сравнению с системой, в которой для передачи сообщений применяются сигналы с QAM без использования помехоустойчивых кодов.

Алгоритм кодирования сообщений

Выше были получены формулы, позволяющие рассчитать количество сигналов $M_{sN}(\Delta\varphi_1)$, входящих в состав АС, а также удельная скорость передачи сообщений в системе связи, использующей ГПФМ модуляцию.

Построенный АС обеспечивает минимальное евклидово расстояние между разными сигналами АС, равное $d \sim \sqrt{N}$, и позволяет передавать по каналу связи различные сообщения, каждому из которых могут быть присвоены номера $m = 0 \dots M_{sN}(\Delta\varphi_1) - 1$. При передаче информационного сообщения, имеющего номер m и представленного в виде информационной двоичной последовательности знаков, эта последовательность поступает на вход системы связи рис. 1.

Для того, чтобы передать такую последовательность по каналу связи с помощью одного из формируемых в модуляторе сигналов $S_m(t)$ с ГПФМ, необходимо преобразовать номер передаваемого сообщения m в последовательность его индексов – чисел I_{mk} , с помощью которых могут быть определены все фазовые координаты CT , соответствующей сигналу $S_m(t)$. Алгоритм решения этой задачи (по Шеннону алгоритм кодирования передаваемого сообщения) должен иметь минимально возможную вычислительную сложность. Предлагаемый рекуррентный алгоритм вычисления чисел I_{mk} (кодирования) основан на разбиении всего множества CT на подмножества 1-го, 2-го, ... $(N-2)$ -го порядка в соответствии со значениями индексов $I_{mi} (i \leq (N-2))$, от которых зависят значения фазовых координат (φ_{mi}) сигнальных точек AC . Последний индекс $I_{m(N-1)}$ определяет координату $\varphi_{m(N-1)}$ у CT , которая лежит на соответствующей окружности и зависит от всех значений φ_{mi} при $i \leq (N-2)$.

Число подмножеств i -го порядка равно $G_i(\vec{I}_{m(i-1)}) = 2M_i(\Delta\varphi_1, \vec{I}_{m(i-1)}) + 1$ – количеству значений, которое могут принимать индексы I_{mi} при определенном значении i , где $M_i(\Delta\varphi_1, \vec{I}_{m(i-1)})$ рассчитываются по формуле (13). Структура множества всех CT для ГПФМ такова, что, например, в каждое из подмножеств 1-го порядка при заданном значении $I_{m1} = const$ входит $G_2(I_{m1}, I_{m2})$ подмножеств 2-го порядка. При этом каждому такому подмножеству относятся CT , первая фазовая координата которых равна $\varphi_{m1} = I_{m1}\Delta\varphi_1$, а вторая $-\varphi_{m2} = I_{m2}\Delta\varphi_2(I_{m1})$. Эта координата определяется значением I_{m2} , изменяющееся в интервале $-M_2(I_{m1}) \leq I_{m2} \leq M_2(I_{m1})$. Подобным же образом в определяются подмножества CT 3-го порядка, координаты φ_{m1} и φ_{m2} которых уже определены (т.е. $I_{m1} = const$ и $I_{m2} = const$). Эти CT входят в $G_3(I_{m1}, I_{m2}, I_{m3})$ подмножеств 3-го порядка и т.д. В подмножества $(N-1)$ -го порядка при $I_{mk} = const$ при $k \leq (N-2)$ входят CT , расположенные на окружностях, причем на каждой из них количество CT равно $M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)})$.

Отметим несколько важных свойств зависимости $M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)})$ от I_{m1} .

1. Значения $M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)})$ при $\vec{I}_{m(N-2)} = const$ зависят, как следует из (15), только от абсолютных значений индексов I_{mk} (т.е. от $|I_{mk}|$). При этом индексы $I_{m(N-1)}$ изменяются в пределах $0 \leq I_{m(N-1)} \leq M_{(N-1)} \times (\Delta\varphi_1, \vec{I}_{m(N-2)})$.

2. Суммарное количество CT , лежащих на всех таких

окружностях, является суммой слагаемых $M_{(N-1)} \times (\Delta\varphi_1, \vec{I}_{m(N-2)m})$ в (16), имеющих одинаковые значения. Число слагаемых, входящих в эту сумму, равно числу возможных комбинаций знаков $+$ и $-$ у индексов $I_{mki} (i = 1, \dots, L)$. Поэтому, если только L значений $|I_{mki}|, |I_{mk2}|, \dots, |I_{mki}|$ не равны нулю, а все остальные равны нулю, то суммарное количество CT в подмножествах $(N-1)$ -го порядка равно $2^L M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)m})$.

3. Номер каждой из 2^L окружностей может быть присвоен, в принципе, произвольно. Но рациональное правило присвоения номера может быть таким: в строке $\vec{I}_{m(N-2)}$, которая входит в формулу для $M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)m})$, на позициях, где расположены индексы $I_{mk} \neq 0$, ставиться число $\gamma_{mk} = 0,5[1 + sign(I_{mk})]$, здесь функция $sign(x) = 1$ при $x > 0$ и $sign(x) = -1$ при $x < 0$, при этом $\gamma_{mk} = 1$ при $I_{mk} > 0$ и $\gamma_{mk} = 0$ при $I_{mk} < 0$. В результате номер соответствующей окружности оказывается записанным в этой строке в двоичной форме.

4. Величины $M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)})$ имеют одно и то же значение при любых перестановках индексов $I_{mk_1}, I_{mk_2}, \dots, I_{mk_L}$.

Для определения индексов I_{mk} необходимо знать количество CT , содержащихся в подмножествах 1-го, 2-го и т.д. порядков. Формула, позволяющая рассчитать количество CT в соответствующем подмножестве k -го порядка в том случае, когда заданы значения индексов $I_{mi} (i \leq (k-1))$:

$$A_k(\vec{I}_{m(k-1)}) = \sum_{I_{mk} = -M_k(\Delta\varphi_1)}^{M_k(\Delta\varphi_1)} \dots \sum_{I_{m(N-1)} = -M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)})}^{M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)})} M_{(N-1)}(\Delta\varphi_1, \vec{I}_{m(N-2)}). \quad (18)$$

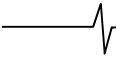
При изменении знака одного из индексов, входящих в строку $\vec{I}_{m(k-1)}$, количество CT в соответствующем подмножестве не изменяется, так как $A_k(I_{m1}, \dots, |I_{m(k-1)}|) = (I_{m1}, \dots, -|I_{m(k-1)}|)$. Отсюда следует, что

$$A_k(|I_{m1}|, \dots, |I_{mi}|, \dots, |I_{m(k-1)}|) = 2A_k(I_{m1}, \dots, I_{m(k-1)}). \quad (19)$$

Абсолютные значения индексов $|I_{mi}|$ могут изменяться в следующих пределах: $0 \leq |I_{mi}| \leq M_i(\Delta\varphi_1, \vec{I}_{m(i-1)})$.

Алгоритм присвоения номеров CT , расположенным на N -мерной сфере, используя (18) и (19) для расчета $A_k(\vec{I}_{m(k-1)})$, таков:

Вычисляют $A_1(|I_{m1}|)$ – количество CT в подмножествах 1-го порядка при разных значениях $|I_{m1}| = 0, \dots, M_1(\Delta\varphi_1)$, а также суммарное количество CT подмножеств 1-го порядка для которых $|I_{m1}| \leq K_{m1} \leq M_1(\Delta\varphi_1)$, равное $A1\Xi(K_{m1}) = \sum_{(|I_{m1}|=0)}^{(K_{m1})} A_1(|I_{m1}|)$. Отме-



тим, что $A1\Xi(M_1(\Delta\varphi_1)) = \sum_{(|I_m|=0)}^{(M_1(\Delta\varphi_1))} A_1(|I_{m1}|) = M_s N(\Delta\varphi_1)$, т.е. суммарное количество CT во всех подмножествах 1-го порядка равно полному количеству сигналов в AC . Поэтому номера всех CT , расположенных на поверхности N -мерной сферы, имеют значения $0 \leq l \leq [M_{sN}(\Delta\varphi_1) - 1]$. Причем, CT , относящиеся к подмножеству 1-го порядка с индексом, равным $|I_{m1}| = 0$, имеют номера $0 \leq l \leq A_1(0) - 1$, если же $|I_{m1}| = 1, \dots, M_1(\Delta\varphi_1)$, то CT этих подмножеств они имеют номера $A_1(|I_{m1}| - 1) \leq l \leq A_1(|I_{m1}|) - 1$.

Вычисляют $M_1(\Delta\varphi_1)$ значений $A1\Xi(K_{m1})$ и, выполняя сравнение числа m с числами $A1\Xi(K_{m1})$, определяют к какому из подмножеств CT 1-го порядка относится номер m . Индекс $|I_{m1}|$ определяют таким образом: если $D_1(K_{m1}) = A_{1\Xi}(K_{m1} - 1) \leq m \leq A_{1\Xi}(K_{m1})$, то $|I_{m1}| = (K_{m1} - 1)$.

Прямое нахождение индекса $|I_{m1}|$ требует выполнения $N_1 = M_1(\Delta\varphi_1)$ сравнений. Однако это число можно существенно уменьшить, если воспользоваться алгоритмом решения задач о взвешивании монет [15]. В [15] показано, что для определения индекса $|I_{m1}|$ достаточно выполнить только $N_{01} = \log_2(N_1)$ сравнений.

Вычислив индекс $|I_{m1}|$ и определив подмножество CT 1-го порядка, содержащее $\Delta N_1 = [A_{1\Xi}(K_{m1}) - A_{1\Xi}(K_{m1} - 1)]$ CT , в котором имеется одна CT , соответствующая номеру m , аналогичным образом вычисляют индекс $|I_{m2}|$ при $|I_{m1}| = const$.

Для этого, найденное подмножество CT 1-го порядка разбивается на $M_2(\Delta\varphi_1, |I_{m1}|)(|I_{m2}| = 0, \dots, M_2(\Delta\varphi_1, |I_{m1}|) - 1)$ подмножеств 2-го порядка. Вычисление индекса $|I_{m2}|$ осуществляется следующим образом:

Вычисляют $A_2(|I_{m1}|, |I_{m2}|)$ – количество CT в подмножествах 2-го порядка при разных значениях $|I_{m2}|$ ($|I_{m2}| = 0, \dots, M_2(\Delta\varphi_1, |I_{m1}|) - 1$), а также суммарное количество CT подмножества 2-го порядка, для которых $|I_{m2}| \leq K_{m2} \leq M_2(\Delta\varphi_1, |I_{m1}|)$, равное $A2\Xi(K_{m2}, |I_{m1}|) = \sum_{(|I_{m2}|=0)}^{(K_{m2})} A_2(|I_{m1}|, |I_{m2}|)$.

Отметим, что $A2\Xi(M_2(\Delta\varphi_1, |I_{m1}|)) = \sum_{(|I_{m2}|=0)}^{(M_2(\Delta\varphi_1, |I_{m1}|))} A_2(|I_{m1}|, |I_{m2}|) = A_1(|I_{m1}|)$, т.е. суммарное количество CT во всех подмножествах 2-го порядка равно полному количеству CT в подмножестве 1-го порядка, которое соответствует индексу $|I_{m1}|$. Причем, CT , относящиеся к подмножеству 2-го порядка с индексом, равным $|I_{m2}| = 0$, CT имеют номера $A_{1\Xi}(K_{m1} - 1) \leq l \leq A_{1\Xi}(K_{m1} - 1) + A_2(0, |I_{m2}|) - 1$, а если $|I_{m2}| = 1, \dots, M_2(\Delta\varphi_1, |I_{m1}|)$, то CT имеют номера $D_1(K_{m1}) + A_2(|I_{m2}| - 1, |I_{m1}|) \leq l \leq D_1(K_{m1}) + A_2(|I_{m2}|, |I_{m1}|) - 1$.

Вычисляют $A_1(|I_{m1}|)$ значений $A2\Xi(K_{m2}, |I_{m1}|)$ и, выполняя сравнение числа m с числами $A2\Xi(K_{m2}, |I_{m1}|)$, определяют к какому подмножеству CT 2-го порядка относится номер m , т.е. определяют индекс $|I_{m2}|$: если $D_2(K_{m2}) = D_1(K_{m1}) + A_{2\Xi}(K_{m2} - 1, |I_{m1}|) \leq m \leq D_1(K_{m1}) + A_{2\Xi}(K_{m2}, |I_{m1}|) - 1$, то $|I_{m2}| = (K_{m2} - 1)$.

Прямой способ нахождения $|I_{m2}|$ требует выполнения $N_2 = M_2(\Delta\varphi_1, |I_{m1}|)$ операций сравнения, а при использовании оптимального метода [11] достаточно выполнить только $N_{02} = \log_2(N_2)$ сравнений. Вычислив индекс $|I_{m2}|$ и определив подмножество CT 2-го порядка, содержащее $\Delta N_2 = [A_{2\Xi}(K_{m2}, |I_{m1}|) - A_{2\Xi}(K_{m2} - 1, |I_{m1}|)]$ CT , в числе которых имеется одна CT , соответствующая номеру m , переходим к определению индекса $|I_{m3}|$.

Индекс $|I_{m3}|$ определяют аналогичным образом, учитывая значения уже найденных индексов $|I_{m1}| = const$ и $|I_{m2}| = const$. Для этого, найденное подмножество CT 2-го порядка разбивается на $M_3(\Delta\varphi_1, |I_{m1}|, |I_{m2}|)(|I_{m3}| = 0, \dots, M_3(\Delta\varphi_1, |I_{m1}|, |I_{m2}|))$ подмножеств 3-го порядка. Продолжая подобные вычисления, последовательно находим индексы $|I_{m1}|, |I_{m2}|, \dots, |I_{m(N-2)}|$.

Следует отметить, что, по мере нахождения этих индексов, интервал чисел, в котором может находиться CT , соответствующая номеру передаваемого m сообщения, все время уменьшается. После определения всех индексов $|I_{m1}|, |I_{m2}|, \dots, |I_{m(N-2)}|$ ($|I_{m(N-2)}| = (K_{(N-2)} - 1)$) он становится равным $\Delta N_{(N-2)} = [A_{(N-2)\Xi}(K_{(N-2)}, \bar{I}_{m(N-2)}) - A_{(N-2)\Xi}(K_{(N-2)} - 1, \bar{I}_{m(N-2)})]$. При вычислении $|I_{mk}|$ находится также нижняя граница этого интервала, которая вычисляется по рекуррентной формуле $D_{(N-2)} \times (K_{m(N-2)}) = D_{(N-3)}(K_{m(N-3)}) + A_{(N-2)\Xi}(K_{(N-2)} - 1, \bar{I}_{m(N-2)})$.

Подмножество CT $(N-1)$ -го порядка, номера которых находятся в интервале $[D_{(N-2)}(K_{m(N-2)}), D_{(N-2)}(K_{m(N-2)}) + \Delta N_{(N-2)}]$. Эти CT лежат на окружностях, число которых, как указывалось выше, равно 2^L (число L равно количеству $|I_{mk}| \neq 0$ в найденной последовательности индексов $\bar{I}_{m(N-2)}$), при этом на каждой окружности лежит $M_{(N-1)}(\Delta\varphi_1, \bar{I}_{m(N-2)m})$ сигнальных точек. Определим число $m_1 = [m - D_{(N-2)}(K_{m(N-2)})]$, которое является относительным номером CT , соответствующей номеру m . Это число определяет относительное положение этой CT , отсчитываемое от первой CT подмножества $(N-1)$ -го порядка и содержащее $Z_{(N-1)}(\bar{I}_{(N-2)m}) = 2^L M_{(N-1)}(\Delta\varphi_1, \bar{I}_{(N-2)m})$ CT .

Порядок нумерации этих окружностей определяется знаками индексов в $\vec{I}_{(N-2)m}$. Номер окружности, на которой расположена CT , соответствующая числу m , равен

$$Y(\vec{I}_{(N-2)m}) = \text{int} \left[\frac{m_1}{M_{(N-1)}(\Delta\varphi_1, \vec{I}_{(N-2)m})} \right].$$

Записав это число в двоичной форме в виде последовательности длиной $(N-2)$ двоичных символов 0 и 1, можно определить знаки в строке абсолютных значений индексов $\vec{I}_{(N-2)m}$, руководствуясь сформулированным выше правилом нумерации окружностей. Если на k -ой позиции двоичной последовательности, определяющей номер окружности, стоит символ 0, то индекс $I_{mk} = -|I_{mk}|$, в противном случае $I_{mk} = |I_{mk}|$. Остаток от деления числа m_1 на $M_{(N-1)}(\Delta\varphi_1, \vec{I}_{(N-2)m})$ определяет значение индекса $I_{m(N-1)}$: если он равен 0, то CT является последней на окружности с номером $Y(\vec{I}_{(N-2)m})$ и $I_{m(N-1)} = M_{(N-1)} \times (\Delta\varphi_1, \vec{I}_{(N-2)m}) - 1$, в противном случае, он равен дробной части результата такого деления, т.е. $I_{m(N-1)} = m_1 - Y(\vec{I}_{(N-2)m})M_{(N-1)}(\Delta\varphi_1, \vec{I}_{(N-2)m})$. Описанный алгоритм позволяет найти все индексы I_{mk} с учетом их знаков.

Процедура нахождения значений I_{mk} , соответствующих номеру m сообщения, поступившего на вход системы связи, требует вычисления K_A чисел $A_i(\vec{I}_{mi}) \times$

$$\times (K_A = \sum_{i=1}^N M_i(\Delta\varphi_1, \vec{I}_{m(i-1)}) \leq N[2M_1(\Delta\varphi_1) + 1])$$

и выполнения $K_S = N \sum_{i=1}^{N-1} \log_2[G_i(\vec{I}_{mi})] \leq N \{ \log_2[2M_1(\Delta\varphi_1) + 1] \}$ сравнений этих чисел с числом m . Таким образом, сложность описанного алгоритма кодирования сообщения растет линейно с увеличением N .

Заключение

В данной работе представлен метод формирования N -мерного ансамбля сигналов (AC) с гиперфазовой модуляцией, для которого все сигнальные точки (CT) соответствующего AC расположены на поверхности N -мерной сферы. Такой AC , как было показано Шенноном [2], является оптимальным при достаточно больших значениях N . Этот AC позволяет обеспечить передачу сигналов в канале связи с ограниченной полосой частот и заданной удельной скоростью передачи и максимально возможной помехоустойчивостью.

Принципиальное отличие этого метода передачи сообщений от тех, которые используются в современных системах связи, состоит в том, что для повышения помехоустойчивости в нем не предусмотрено внесение в последовательность информационных символов передаваемого сообщения помехоустойчивых кодов, содержащих помимо информационных дополнительные избыточные (проверочные) символы. Внесение избыточных

символов снижает скорость передачи сообщений, хотя и позволяет на приемном конце линии связи исправлять ошибки в информационной последовательности, возникшие при демодуляции принятого сигнала.

Описан алгоритм гиперфазовой модуляции – метода формирования сигналов $S_m(t)$ такого ансамбля, а также разработан эффективный в вычислительном отношении метод кодирования сообщений – процедуры вычисления по номеру m передаваемого сообщения индексов модуляции I_{mk} , которые определяют фазовые координаты CT , соответствующей сигналу $S_m(t)$. Применение рассмотренного оптимального N -мерного AC даже при сравнительно небольших значениях N реализует выигрыш в величине минимального расстояния между CT в оптимальном AC по сравнению с широко распространенным AC с QAM . С увеличением N этот выигрыш растет.

Литература

1. Shannon C. Communication in the presence of noise, Proc. IRE, № 1, 1949. (Перевод на русский язык статьи «Связь при наличии шума», опубликованной в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы / под ред. Р.Л. Добрушина и О.Б. Лупанова).
2. Shannon C. Probability of error for optimal codes in Gaussian channel. Bell System Techn. J., May, 1959. (Перевод на русский язык статьи «Вероятность ошибки для оптимальных кодов в гауссовском канале», опубликованной в книге Шеннон К. Работы по теории информации и кибернетике. М.: Из-во иностранной литературы. // Под ред. Р.Л. Добрушина и О.Б. Лупанова).
3. Weinstein S.B., Ebert P.M. Data transmission by frequency division multiplexing using the discrete Fourier transform. IEEE Trans. COM-19, № 10, 1971.
4. Быховский М.А. Вероятность ошибки для оптимальных многомерных кодов в гауссовом канале связи и их основные характеристики. Электросвязь № 2, 2016.
5. Быховский М.А. Помехоустойчивость приема оптимальных сигналов, расположенных на поверхности N -мерного шара. Электросвязь № 3, 2016.
6. Прокис Дж. Цифровая связь. //Перевод с английского под ред. Д.Д. Кловского. //М.: Советское радио, 2000.
7. Дж. Кларк мл., Дж. Кейн. Кодирование с исправлением ошибок в системах цифровой связи. М.: Радио и связь, 1987
8. Ungerboeck G. Trellis-Coded Modulation with Redundant Signal Sets. PartlandII, IEEE Commun. Mag., vol. 25, № 2, 1987.
9. Fomey G.D., Gallager R.G., Lang G.R., Longstaff F.M., Qureshi S.U. Efficient Modulation for Bandlimited Channels. IEEE J. Select Areas in Commun. vol. SAC-2, n. 5, № 9, 1984.
10. Fomey G.D., Wei L.F. Multidimensional constellations-Part I: Introduction, figures of merit and generalized cross constellations. IEEE I. Select. Areas Commun., vol. 7, № 8, 1989.
11. Fomey G.D. Multidimensional constellations-Part II: Voronoi constellations. IEEE I. Select. Areas Commun., vol. 7, № 8, 1989.
12. Бураченко Д.Л., Савищенко Н.В. Геометрические модели сигнально-кодовых конструкций. СПб.: Военная академия связи, 2012.
13. Галлагер Р. Коды с малой плотностью проверок на четность. – М.: Мир, 1966
14. N -мерная сфера. <https://en.wikipedia.org/wiki/N-sphere>.
15. Альсведе Р., Вергенер И. Задачи поиска. М.: Мир, 1982.