В. Д. Гоппа, Коды, ассоциированные с дивизорами, *Пробл. передачи информ.*, 1977, том 13, выпуск 1, 33–39

ПРОБЛЕМЫ ПЕРЕДАЧИ ИНФОРМАЦИИ

Том XIII 1977 Вып. 1

УДК 621.391.15

КОДЫ, АССОЦИИРОВАННЫЕ С ДИВИЗОРАМИ *

В. Д. Гоппа

Вводится понятие линейного кода, ассоциированного с дивизором в поле алгебраических функций. Мощность и вес таких кодов оцениваются с помощью теоремы Римана – Роха. Циклические коды и (L, g)-коды являются частным случаем рассматриваемых кодов.

§ 1. Теорема Римана — Роха

Пусть k — произвольное поле. *Полем K алгебраических функций* одной переменной над k называется конечно-порожденное расширение поля k степени трансцендентности 1 [1, 2]. Таким образом, алгебраическая функция y есть корень уравнения $F(x, y)=0$, коэффициентами которого являются рациональные функции от x. С другой стороны, уравнение $F(x, y)=0$ можно рассматривать как уравнение кривой на плоскости с координатами x и y. Поле K совпадает с полем рациональных функций на плоской кривой C. С полем K естественным образом связана также проективная кривая X (не обязательно плоская) — *неособенная модель* K,— точки которой находятся во взаимно-однозначном соответствии с кольцами дискретного нормирования поля K.

Кольцом дискретного нормирования называется кольцо R, все необратимые элементы которого образуют идеал M, являющийся единственным максимальным идеалом в R. Идеал M является главным: $M=(t)$, а любая образующая этого идеала называется *униформизирующим параметром*. Каждый элемент z кольца R представляется однозначно в виде $z=ut^n$, где u — единица кольца R, а n — неотрицательное целое.

Кольца дискретного нормирования поля K обычно называют *точками поля K*. Каждая точка порождает нормирование поля, связанное с единственным представлением алгебраической функции $f \in K$ в форме $f=ut^n$, где u — единица кольца R, а n — целое. Величина n не зависит от конкретного выбора униформизирующего параметра t, называется *порядком функции в точке P* и обозначается $n=\mathrm{ord}_P(f)$. Обычно предполагают $\mathrm{ord}_P(0)=\infty$ в любой точке P.

С геометрической точки зрения кривые C и X связаны бирациональным морфизмом φ: $X \to C$, а кольцо R совпадает с локальным кольцом в точке P, т. е. с множеством рациональных функций на X, регулярных (определенных) в точке P. Если $n=\mathrm{ord}_P(f)>0$, то говорят, что P является *нулем кратности n* функции f, а при $n<0$ — *полюсом кратности* $(-n)$. Элементы поля K, алгебраические над k, называются *константами*. Они образуют подполе поля K, называемое *полем констант*. Фактор-кольцо R/M является полем, которое называется *полем вычетов в точке P*. Это поле имеет конечную степень над полем констант, которая называется *степенью точки P* и обозначается $d(P)$.

* К 25-летию выхода в свет статьи Р. Хэмминга «Коды для обнаружения и исправления ошибок».

Дивизором поля K называется формальная сумма

$$D=\sum_{P\in X} n_P P, \quad n_P\in Z$$

и $n_P=0$ для всех точек P, кроме конечного числа. Все дивизоры поля K образуют абелеву группу — свободную абелеву группу на множестве X. *Степенью дивизора* называется величина $\deg(D)=\sum_{P\in X} n_P d(P)$. Дивизор считают *положительным* (*эффективным*), если все $n_P\geqslant 0$ и пишут $\sum n_P P\geqslant\sum m_P P$, если все $n_P\geqslant m_P$.

Пусть $f\in K$. *Дивизором функции* $f\neq 0$ называется дивизор

$$\operatorname{div}(f)=\sum_{P\in X}\operatorname{ord}_P(f)P.$$

Это определение корректно, поскольку любая алгебраическая функция имеет конечное число нулей и полюсов. Для любой функции справедливо равенство $\deg(\operatorname{div}(f))=0$, которое означает, что число нулей равно числу полюсов (с учетом кратности и степени).

Пусть $D=\sum n_P P$ — дивизор на X, L^+ и L^- множества точек, для которых $n_P>0$ и $n_P<0$ соответственно. Если для некоторой функции $f\in K$ справедливо неравенство $\operatorname{div}(f)\geqslant -D$, т. е. $\operatorname{ord}_P(f)\geqslant -n_P$ для всех P, то это означает: 1) функция f не имеет полюсов вне множества L^+; 2) кратность каждого полюса не превышает n_P для $P\in L^+$; 3) кратность каждого нуля в точках множества L^- не меньше $(-n_P)$.

Все такие функции образуют конечномерное векторное пространство, которое называется *пространством, ассоциированным с дивизором* D и обозначается $L(D)$. Вычисление размерности этого пространства составляет содержание *проблематики Римана — Роха*. Знаменитая теорема Римана — Роха утверждает, что

$$l(D)=\deg(D)-g+1+l(W-D),$$

где $l(D)=\dim L(D)$, g — род поля K, а W — канонический дивизор (представитель класса эквивалентности дивизоров дифференциалов поля K). Если $\deg(D)\geqslant 2g-1$, то эта теорема совпадает с *оценкой Римана*: $l(D)=\deg(D)+1-g$.

§ 2. Коды, исправляющие ошибки

Пусть E — конечное множество (алфавит), E^n — множество наборов (слов) из n букв алфавита E. Для любых двух слов $x, y\in E^n$ *расстояние по Хэммингу* $d(x, y)$ определяется как число символов, в которых они отличаются. Кодом называется любое подмножество $U\subseteq E^n$, число элементов U называется *мощностью кода*, а величина $d(U)=\min_{x,y\in U} d(x, y)$, $x\neq y$ — *кодовым расстоянием*. Если $d(U)=2t+1$, то *код исправляет* t *ошибок*. Требуется при фиксированных n и t найти код максимальной мощности.

Такая проблема была сформулирована Хэммингом в статье [3], положившей начало теории корректирующих кодов. Сам Хэмминг полностью решил проблему для случая $t=1$. *Коды Хэмминга*, представляющие собой высокий образец изобретательности, получили широкое распространение

в технике, а использованный им метод проверок на четность оказал решающее влияние на все дальнейшее развитие теории.

Рожденная в середине 20-го века, новая теория «генетически» восприняла основные тенденции современной математики, и с самых первых дней своего существования начала стремительное движение к алгебраизации. Уже в первое десятилетие термины «группа», «кольцо», «идеал», «поле Галуа» заняли прочное место в литературе по теории кодов [3–11].

Понятие *линейного кода*, выделенное Слепяном [7], включает в себя следующие элементы: алфавит E наделяют структурой конечного поля $k=GF(q)$, E^n трактуют как векторное пространство над k с фиксированным базисом, а в качестве кодов рассматривают подпространства E^n. Число ненулевых координат в разложении вектора x называют *весом Хэмминга* $d(x)$. Вес Хэмминга является нормой на E^n. Для линейного кода $d(x, y)=d(x-y)$,

$$d(U)=\min_{x\in U} d(x), \quad x\neq 0.$$

Декодирование для линейного кода заключается в нахождении вектора минимального веса в смежном классе по этому коду. Для того чтобы код мог быть применен в реальной аппаратуре, он должен допускать простые процедуры кодирования и декодирования. В то время как кодирование для любого линейного кода не вызывает трудностей (в связи с тем что линейное пространство допускает экономное представление с помощью базиса), декодирование, отличное от перебора, является значительно более трудной задачей и эффективные алгоритмы найдены лишь для некоторых специальных подклассов линейных кодов.

Применение корректирующих кодов особенно выгодно при больших длинах n. В 1952 г. Гилбертом [12] была получена асимптотическая граница, указывающая на потенциальную осуществимость некоторого соотношения между избыточностью кода и его корректирующей способностью. Самый приятный комплимент, который можно сделать в настоящее время любому достаточно длинному коду,— сказать, что он достигает границы Гилберта. Многие специалисты склоняются к мысли, что эта граница вообще не может быть улучшена.

Переход от произвольных кодов к линейным происходит без потери лучших длинных кодов в том смысле, что уже среди линейных кодов существуют коды, достигающие границы Гилберта. Этот важный факт был установлен Варшамовым [13] в 1957 г.

Если ограничиться только линейными кодами, то теория корректирующих кодов становится разделом линейной алгебры над конечным полем, а специфика теории связана с весом Хэмминга: если задано произвольное линейное пространство, то основной вопрос, который интересует специалиста по теории кодов,— сколько ненулевых координат имеют векторы этого пространства в разложении по некоторому базису. В матричной терминологии эта специфика может быть выражена в следующей эквивалентной форме: каково то наибольшее число d (вес матрицы), что все миноры порядка $<d$ не равны нулю?

Как ни странно, единственная известная область физики и математики, где возникал ранее в связи с матрицами подобный вопрос,— это теория малых колебаний механических систем. Осцилляционные матрицы, интенсивно изучавшиеся Гантмахером и Крейном [14, 15], определены, правда, над полем вещественных чисел, а не над конечным полем, и к ним предъявляется более сильное требование — все миноры порядка $<d$ должны быть положительными. Тем не менее две самые известные осцилляционные матрицы — матрица Вандермонда и матрица Коши — играют важную роль в современной теории кодов.

§ 3. Коды, ассоциированные с дивизорами

Пусть k — конечное поле $GF(q)$, k^* — его расширение $GF(q^m)$, $D=\sum n_P P$ и $G=\sum m_Q Q$ — дивизоры поля алгебраических функций над k^*, причем G — эффективный дивизор и все P, Q различны. Отождествим пространство $L(D)$ с множеством слов E^n, а подпространство $L(D-G)\subseteq L(D)$ назовем *кодом, ассоциированным с дивизором* D, или (D, G)-*кодом*. Точнее говоря, в пространстве $L(D)$ фиксируем некоторый базис и в качестве q-ичного кода выбираем линейные комбинации базисных функций с коэффициентами из поля $k=GF(q)$, которые входят в пространство $L(D-G)$.

Мощность и длина таких кодов определяются с помощью теоремы Римана — Роха и не зависят от выбора базиса. Для значения же кодового расстояния выбор базиса играет решающую роль, и теорема Римана — Роха подсказывает один из удачных вариантов выбора этого базиса. Действительно, составим базис из функций с наименьшей степенью дивизора полюсов. Если $f\in L(D-G)$, то эта функция имеет большое число нулей (их количество зависит, в частности, от степени дивизора G). Но тогда f имеет и большое число полюсов, а поскольку базисные функции имеют мало полюсов, то кодовая функция f должна иметь много ненулевых координат в разложении по этому базису.

Рассмотрим два примера реализации этой идеи в вырожденном случае поля рода 0. Поле алгебраических функций совпадает в данном случае с полем $k^*(z)$ рациональных функций над k^*. Все точки этого поля легко перечисляются: каждому неприводимому над k^* многочлену p ставится в соответствие точка P, которой соответствует кольцо дискретного нормирования R_P, состоящее из всех дробей вида g/h, где h не делится на p. Различные неприводимые многочлены приводят к различным точкам. Наконец, существует еще одна точка, которая обозначается ∞. Кольцо этой точки состоит из всех дробей g/h, у которых $\deg g\leqslant\deg h$.

Униформизирующим параметром конечной точки P является соответствующий неприводимый многочлен $p(z)$, а $\mathrm{ord}_P(f)$ в этой точке равен целому n в представлении $f=p^n g/h$, где g и h не делятся на p. В бесконечной точке униформизирующим параметром является функция $1/z$, и если $f=g/h$, то $\mathrm{ord}_\infty(f)=\deg h-\deg g$.

1) Пусть L — некоторое множество точек первой степени поля $k^*(z)$, т. е. $L=\{z-\alpha_1, z-\alpha_2, \ldots, z-\alpha_n\}$, где все $\alpha_j\in k^*$, $\alpha_i\neq\alpha_j$. Выберем многочлен $g(z)=p_1^{m_1}p_2^{m_2}\ldots p_k^{m_k}$ над полем k^*, где все p_i неприводимы над k^* и не принадлежат множеству L. Положим $D=\sum_{P\in L} P-\infty$. Пространство $L(D)$ состоит в данном случае из всех дробей вида $\sum\gamma_i/(z-\alpha_i)$, $\gamma_i\in k^*$. Определим далее эффективный дивизор $G=\sum m_i Q_i$, где точка Q_i соответствует неприводимому многочлену p_i в разложении $g(z)$. Пространство $L(D-G)$ высекает из $L(D)$ те дроби, числители которых делятся на $g(z)$. Таким образом, (D, G)-код в данном случае совпадает с (L, g)-*кодом* [16].

Так как $l(D)=\deg(D)+1-g$, а $\deg(D)$ в этом примере равна $n-1$, то $l(D)=n$. Поскольку $\deg(G)=\sum m_i\deg p_i=\deg g(z)$, заключаем, что размерность кода $L(D-G)$ над k^* равна $n-\deg g(z)$, а над полем k — не меньше величины $n-m\deg g(z)$. В качестве базиса можно выбрать в данном случае все дроби $1/(z-\alpha_i)$. Так как любая ненулевая функция из $L(D-G)$ имеет дивизор нулей степени $\geqslant\deg g(z)+1$, то она имеет по крайней мере $\deg g(z)+1$ ненулевых координат в разложении по выбранному базису.

2) Выберем теперь дивизор $D=(n-1)\infty$. Пространство $L(D)$ состоит в данном случае из всех многочленов над k^* степени $\leqslant n-1$, а $l(D)=n$. Любой эффективный дивизор $G=\sum m_Q Q$, где Q не равно ∞, порождает (D, G)-код, который совпадает с циклическим кодом [8]. Отметим, что выбор в качестве базиса функций $\{1, z, z^2, \ldots, z^{n-1}\}$ не позволяет в данном случае оценить вес кодовых слов.

Циклические коды были введены Прейнджем в 1957 г. Эти коды допускают очень простую реализацию и создали своеобразную «техническую эстетику» кодов, основанную на широком использовании регистров сдвига с обратными связями. Такие регистры, применяемые для кодирования, обнаружения и исправления ошибок, шифрования сообщений в целях секретности, вычислений в полях Галуа и для получения высококачественных псевдослучайных последовательностей, являются превосходным «универсальным» средством в современной технике связи.

(L, g)-коды появились в 1970 г. в результате случайно обнаруженной связи теории кодов с теорией колебаний. Они включают в себя в качестве подклассов полученные ранее *БЧХ-коды* [9, 10], *коды Сриваставы* [17] и *коды Габидулина* [18].

БЧХ-коды являются единственными кодами, лежащими в пересечении циклических кодов с (L, g)-кодами. Открытые Хоквингемом в 1959 г., переоткрытые затем Боузом и Рой — Чоудхури, эти коды составили «главное направление» в теории кодов. БЧХ-коды включают в себя в качестве подклассов коды Хэмминга [3], *код Голея* [4] и *коды Рида — Соломона* [11].

Долгое время несколько изолированное положение в теории занимали коды Рида — Маллера. В отличие от других кодов, для которых сначала удавалось сравнительно просто оценить кодовое расстояние, а затем предпринимались попытки реализовать это расстояние, в случае кодов Рида — Маллера сразу же был найден замечательный алгоритм исправления многократных ошибок, но природа высокой корректирующей способности этих кодов оставалась неясной до 1968 года, когда Касами, Лин и Питерсон [19] обнаружили, что код Рида — Маллера любого порядка получается из БЧХ-кода удалением определенного числа векторов. Таким образом, оценка веса этих кодов, так же как и БЧХ-кодов и всех (L, g)-кодов, — все та же оценка Римана, хотя и в несколько тривиальной ситуации.

Переход от произвольных линейных кодов к кодам, ассоциированным с дивизором, совершается также без потери лучших длинных кодов — уже среди неприводимых (L, g)-кодов встречаются коды, достигающие границы Гилберта [16]. Существует надежда, что с помощью сильных методов алгебраической геометрии удастся расклассифицировать эти коды, т. е. отделить «хорошие» дивизоры от «плохих». Но более заманчивым кажется переход к «настоящим» алгебраическим функциям, т. е. к полям ненулевого рода, а, может быть, и к многообразиям высшей размерности.

Задача декодирования для (D, G)-кода сводится к задаче нахождения функции $f \in L(D)$, имеющей минимальное число ненулевых координат в разложении по базису и такой, что $(f-S) \in L(D-G)$. Здесь S — функция из $L(D)$, соответствующая вектору на выходе канала связи (синдром). В случае (L, g)-кодов эта задача эквивалентна нахождению дробно-рациональной функции ψ/φ минимальной степени, приближающей формальный ряд $S(z)$ в g-адической топологии:

(a) $$S-\psi/\varphi \equiv 0 \pmod g.$$

Частный случай $g=z^{2t}$ соответствует БЧХ-кодам:

(aa) $$\psi - S\varphi \equiv 0 \pmod{z^{2t}}.$$

Первый и самый простой подход к решению указанного сравнения заключается в составлении системы линейных уравнений относительно неизвестных коэффициентов многочленов ψ и φ. На этом пути появляются матрицы Ганкеля [14], а в теории кодов это соответствует алгоритму Питерсона [20]. Алгоритм Питерсона был получен, когда существовала еще «наивная» точка зрения на БЧХ-коды, связанная с матрицей Вандермонда, и когда делались лишь первые шаги на пути построения алгебраических схем декодирования. Введенное им понятие многочлена-локатора ошибок стало одним из центральных понятий теории.

Берлекэмп [21] свел задачу декодирования БЧХ-кодов к решению сравнения (аа), которое он назвал «ключевым уравнением» для БЧХ-кодов. Для решения этого сравнения им был построен эффективный алгоритм, основанный на переходе от сравнения по модулю z^i к сравнению по модулю z^{i+1}. Таким образом, Берлекэмп пришел по существу к идее подъема в m-адических топологиях.

Третий метод декодирования [22], несколько уступающий по простоте реализации алгоритму Берлекэмпа, был получен в результате манипуляции терминологией: (а) может быть переписано в виде

$$\|S-\psi/\varphi\|_g \leqslant \Theta, \quad \|\varphi\|_\infty \text{ минимальна},$$

где $\|\ \|_g$ обозначает g-адическую псевдонорму, а $\|\ \|_\infty$ — норму в точке ∞. В таком виде задача декодирования (L, g)-кодов вполне эквивалентна той задаче, с которой столкнулся в 17-м веке Гюйгенс при построении модели солнечной системы с помощью зубчатых колес [23], точнее, g-адическому варианту этой задачи, сформулированному Малером [24]. Алгоритм цепных дробей, примененный Малером для наилучшего рационального приближения g-адических чисел, переносится без изменений на функциональные поля.

Продолжая знакомство с литературой, в которой встречается слово «нормирование», автор убедился, что все рассуждения о нормированиях в функциональных полях ведут к теореме Римана — Роха. Так возникло впечатление неизбежности встречи теории кодов с центральной проблематикой алгебраической геометрии.

ЛИТЕРАТУРА

1. *Шафаревич И. Р.* Основы алгебраической геометрии. М., «Наука», 1972.
2. *Chevalley C.* Introduction to the theory of algebraic functions of one variable. New York, Amer. Math. Soc., 1951.
3. *Hamming R.* Error Detecting and Error Correcting Codes. Bell System Techn. J., 1950, 29, *2*, 147–160.
4. *Golay M.* Notes on Digital Coding. Proc. IRE 1949, 37, *6*, 657.
5. *Muller D.* Application of Boolean Algebra to Switching Circuit Design and to Error Detection. IEEE Trans. Electr. Comp., 1954, 3, *1*, 6–12.
6. *Reed I.* A Class of Multiple-Error-Correcting Codes and the Decoding Scheme. IEEE Trans. Inform. Theory, 1954, 4, *1*, 38–49.
7. *Slepian D.* A Class of Binary Signaling Alphabets. Bell System Techn. J., 1956, 35, *1*, 203–234.
8. *Prange E.* Cyclic error-correcting codes in two symbols. Air Force Cambridge Res. Center Techn. Notes, 1957, 57–103.
9. *Hocquenghem A.* Codes correcteurs d'erreurs. Chiffres, 1959, 2, *2*, 147—156.
10. *Bose R., Ray-Chaudhuri D.* On a Class of Error-Correcting Binary Group Codes. Inform. and Control, 1960, 3, *1*, 68–79.
11. *Reed I., Solomon G.* Polynomial Codes over Certain Finite Fields. J. Soc. Industr. Appl. Math., 1960, 8, *2*, 300–304.
12. *Gilbert E.* A Comparison of Signaling Alphabets. Bell System Techn J., 1952, 31, *10*, 504–522.
13. *Варшамов Р. Р.* Оценка числа сигналов в кодах с коррекцией ошибок. Докл. АН СССР, 1957, 117, *5*, 739–741.
14. *Гантмахер Ф. Р.* Теория матриц. М., «Наука», 1967.

15. *Гантмахер Ф. Р., Крейн М. Г.* Осцилляционные матрицы и ядра и малые колебания механических систем. М., Гостехиздат, 1950.
16. *Гоппа В. Д.* Рациональное представление кодов и (L, g)-коды. Проблемы передачи информации, 1971, 7, *3*, 41–49.
17. *Srivastava J.* Unpublished remarks at the combinatorial symposium. University of North Carolina. N. G., Chapel Hill, 1967.
18. *Габидулин Э. М.* Декодирование максимальных кодов. Тр. 3-й конф. по теории передачи и кодирования информации. Москва – Ужгород, 1967, 51–62.
19. *Kasami T., Lin S., Peterson W.* New Generalisations of the Reed — Muller Codes. IEEE Trans. Inform. Theory, 1968, 14, *2*, 189–199.
20. *Peterson W.* Error-correcting codes. M. I. T. Press, Cambridge, Mass., 1961.
21. *Berlekamp E.* Algebraic coding theory. New York, McGraw-Hill, 1968.
22. *Гоппа В. Д.* О декодировании (L, g)-кодов. Докл. АН СССР, 1975, 222, *6*, 1309–1310.
23. *Хинчин А. Я.* Цепные дроби. М., Физматгиз, 1961.
24. *Mahler K.* Lectures on diophantine approximations. University of Notre-Dome, 1961.

Поступила в редакцию
4 июня 1975 г.