

Предсказание некоторых случайных процессов

А.В. Крамаренко

Много лет назад в город Харьков, который тогда был крупным научным и промышленным центром, пришли свобода и демократия вместе с цивилизацией. На железнодорожном вокзале поставили игральные автоматы – примитивных одноруких бандитов. Естественно, ещё непорочный в своей наивности народ проигрывал все наличные деньги, включая те, которые были необходимы для оплаты проезда. Успех владельцев автоматов был чрезвычайным: деньги пассажиров текли рекой.

Однако через очень короткое время проигрывать стали автоматы, а их владельцы нести убытки. Это был беспорядок - проигрывать по определению всегда должен народ, и к делу подключилась полиция. Довольно быстро выяснилось, что если сделать простой прибор с такой же как у однорукого бандита микросхемой генератора случайных чисел, то после нескольких проигрышей можно синхронизировать генераторы, а дальше игра пойдёт с точным предсказанием, и у бандита нет шансов. «Злоумышленников» с генератором, естественно, поймали и посадили в тюрьму, хотя они всего лишь делали то же, что и миллионы игроков на бирже, которых в тюрьму, тем не менее, не сажают.

Разумеется, этот пример не имеет отношения к действительному анализу случайного процесса и приведен всего лишь для иллюстрации выгод эффективного предсказания.

Первым в мире реальным работающим предсказателем был прибор, сконструированный великим Уильямом Томсоном (лордом Кельвином). В этом анализаторе механическое (!) вычисление гармоник Фурье многие годы обеспечивало точное предсказание уровня приливов на побережье США.

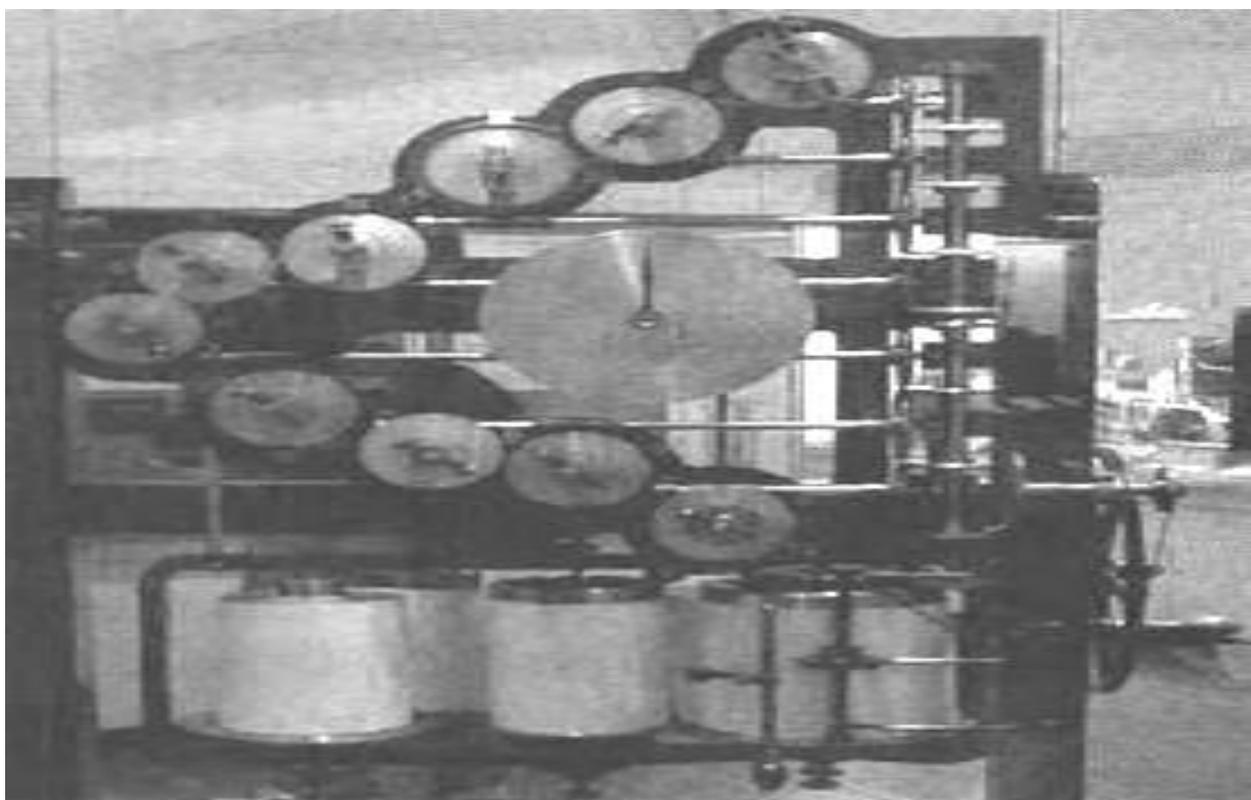


Рис.1. Гармонический анализатор приливов, выставленный в Музее Науки в Лондоне.

Приливы, впрочем, заведомо не вполне случайный процесс, и автор явно пытался развить метод. «В своих лекциях Кельвин часто показывал слушателям причудливые конструкции из различных элементов (стальных кабелей, маятников, деревянных решеток с гириями на конце, маховиков, брусков, пружин и так далее). Индивидуальное поведение каждого элемента было хорошо известно, но все вместе они порождали бесконечное число состояний движения, которые было довольно сложно рассчитать» [1].

Заметим, что это был именно генератор случайного потенциально предсказуемого процесса. Кельвин явно пытался найти решение, для которого у него не было достаточно мощной вычислительной техники.

Современное положение с анализом и предсказанием случайных и не вполне случайных процессов очень разнообразно, если допустимо так выразиться: встречаются как сложнейшие работы с серьезной и трудно понимаемой математикой, так и попытки реанимировать циклы Кондратьева, Слуцкого и подобную чепуху. Результаты применения новых и старых методов тоже различны, предсказания таких процессов как, например, экономическая активность, пока не слишком успешны.

Много лет назад реаниматологи в ЦКБ-5 г. Харькова пожелали иметь для выбора медикаментов в процессе интенсивной терапии график какого-то физиологического параметра, который отображает процесс «вперёд». Мысль явно опередила время и была совершенно верной – это прогноз «на кончике иглы», причём планировалось, что результаты будут поступать в реальном времени, и это позволит своевременно корректировать терапию.

Таким параметром оказалось красное смещение в энергетическом спектре PQRST комплексов электрокардиограммы. Т.е. по мере рассогласования генератора (мышцы сердца) с антенно-фидерным устройством (сосудистой системой) возрастает КСВ (коэффициент стоячей волны), увеличивается нагрузка на сердечную мышцу, и ухудшается кровоснабжение: при этом энергетический спектр ЭКГ комплекса смещается в красную область (визуально эффект почти не заметен). Разумеется, введением медикаментозных препаратов и коррекцией электролитного баланса реаниматологи стремятся компенсировать развивающуюся недостаточность кровоснабжения. И если прогноз по красному смещению изменяется после введения препарата, то поправки в тактику лечения могут быть сделаны не только своевременно, но даже с упреждением, что для реанимации исключительно важно. Но для работы с упреждением необходимо именно предсказание поведения кривой красного смещения. Особая трудность заключается в том, что эта кривая изменяется как под влиянием физиологических процессов, так и под влиянием патологических изменений и, естественно, медикаментозной терапии.

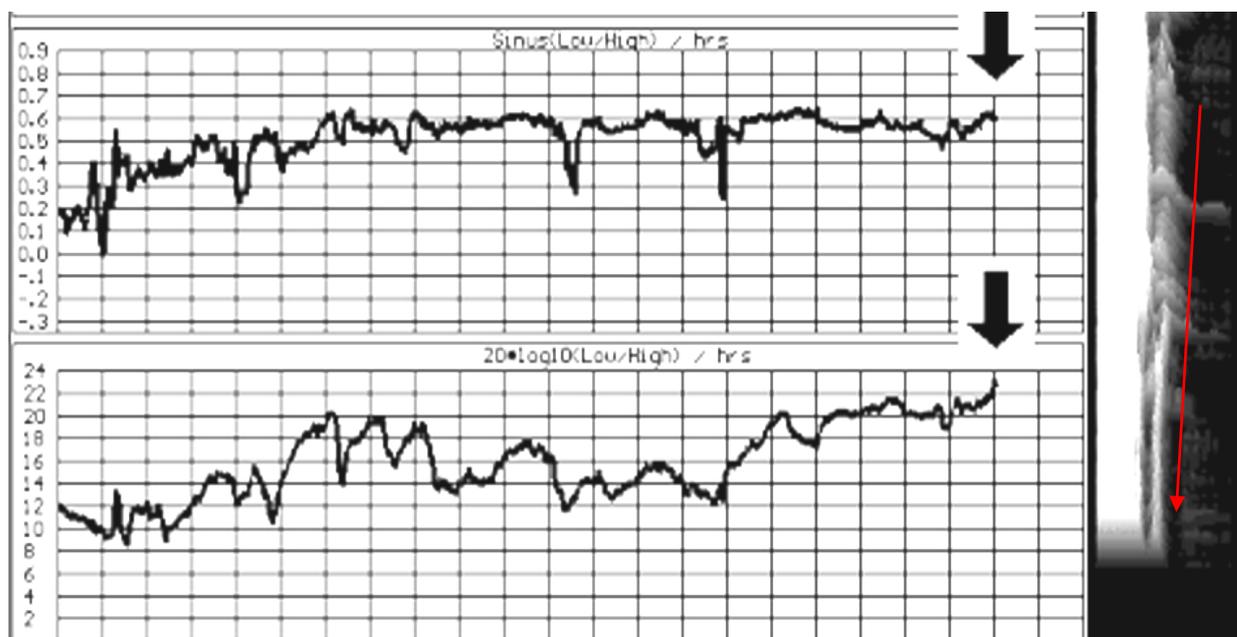


Рис.2. График динамики красного смещения. Острый трансмуральный переднеперегородочно-верхушечный инфаркт миокарда с образованием острой аневризмы с пристеночным тромбом. Причина смерти (этот момент обозначен стрелкой) - линейный разрыв передней стенки левого желудочка с гемотампонадой сердца. Наглядно видно как на спектрограмме справа, (ранние спектры выше поздних) исчезают высокочастотные компоненты.

Этот проект после трудоемких теоретических и экспериментальных работ всё-таки состоялся, и результаты оказались даже лучше ожидаемых. Самое главное - метод

одобрили реаниматологи. Развития, впрочем, работа не получила: энтузиасты уехали в более перспективные страны, а государственным структурам это было неинтересно.

В общем – обычная история интеллектуальной инволюции. Следы остались только в виде статьи [2].

И все же позволю себе привести только один из многих и многих примеров предсказаний, в которые почти невозможно поверить, а естественный скепсис появляется даже у самых доброжелательно настроенных коллег.

Иллюстрации показывают результаты прогнозирования величины красного смещения. Системе известно поведение графика только до (NB!) серой вертикальной линии на рисунке.

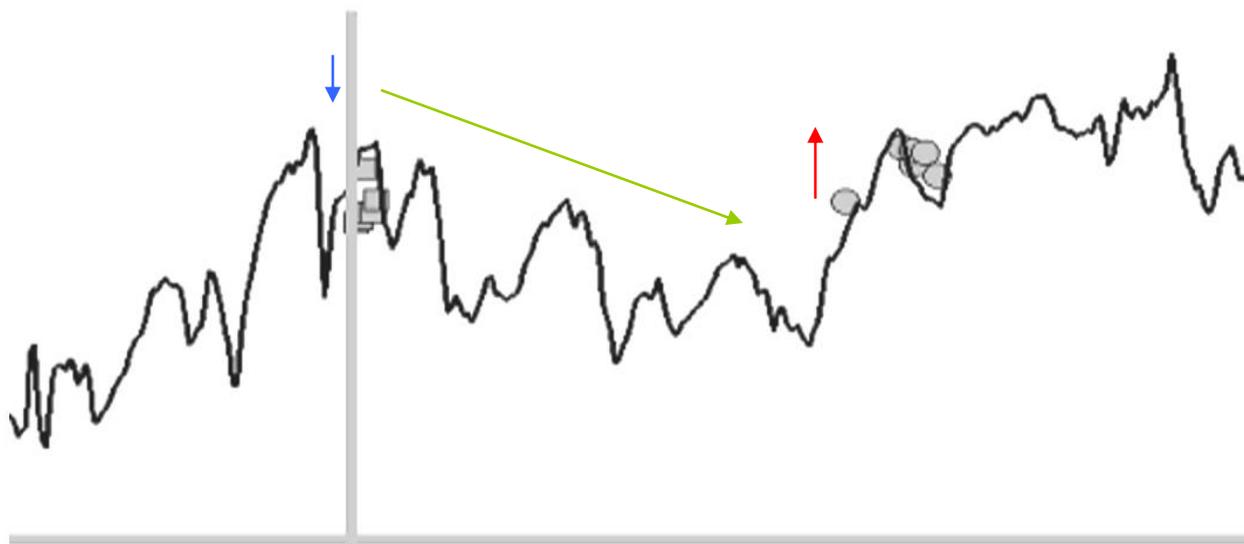


Рис. 3. Тот же случай тяжелой кардиальной патологии, 22 часа наблюдения. Квадратными метками обозначается кратковременный (на 30 минут) прогноз, долговременный (на 12 часов) – круглыми. В момент, обозначенный синей стрелкой, начата дополнительная медикаментозная терапия: клинически – наступило существенное улучшение (зеленая стрелка), но долговременный прогноз не подтверждает этого.

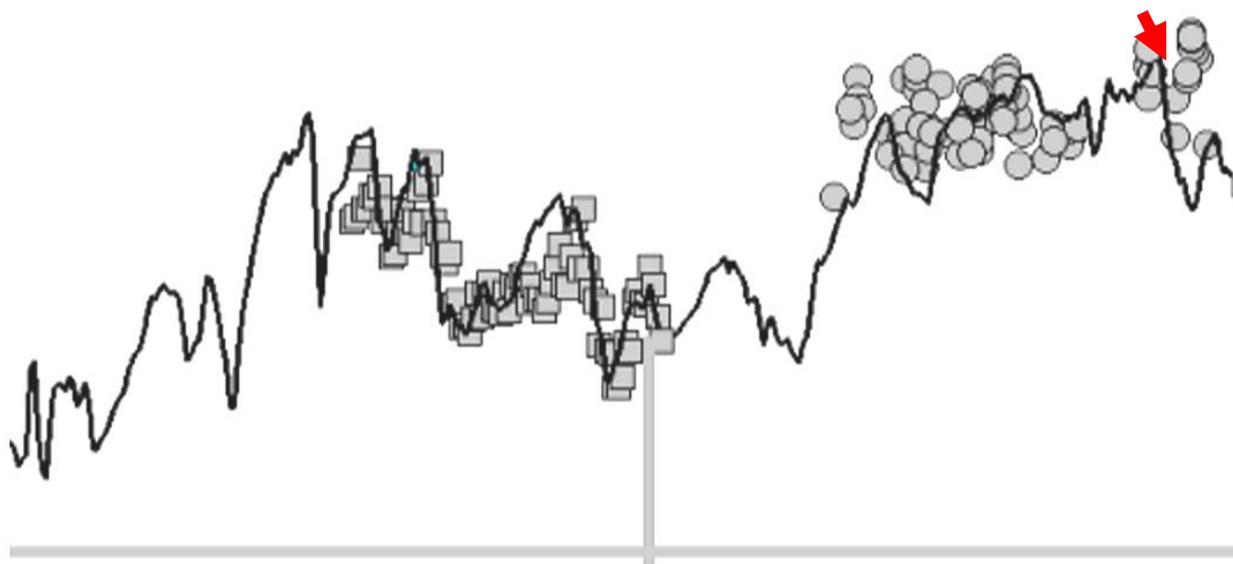


Рис.4. Ситуация через 7 часов. По-прежнему системе известна кривая только до (NB!) серой вертикальной линии. Красной стрелкой показан предсказанный критический подъем графика, в то время как по всем клиническим и аппаратным данным состояние больного заметно улучшалось под влиянием терапии. Катастрофа была предсказана за 12 часов до ее проявления [3].

Можно, разумеется, предположить, что в представленных случаях имеет место детерминированный процесс, и предсказания не представляют никакой сложности. Но если есть явная реакция кривой на введение различных медикаментозных препаратов, то инвариантность процесса исключается (увы, для того, чтобы понять это, потребовалось немало времени).

Но оставим изувеченную ковидом и парацетамолом медицину и вернемся к «причудливым конструкциям из различных элементов (стальных кабелей, маятников...)». Сделаем допущение, что при устремлении количества «кабелей и маятников» к очень большому числу, а в пределе – бесконечности, можно получить не вполне полигармонический процесс [4], который ничем не отличается от истинного шума. Т.е. этот сигнал может иметь даже равномерный спектр и нормальное распределение Гаусса, но будет предсказуемым шумом Кельвина (думается, такое название будет уместно из уважения к великому ученому).

В сущности, сумма большого количества нестабильностей Кельвина-Тейлора и будет, например, таким шумом [5]. Кельвиновские шумы также обязаны существовать в рамках предложенной Греем Уолтером гипотезы: «стабилизация через флуктуации», которая уверенно подтвердилась в нейрофизиологии. (Альфа-ритм электроэнцефалограммы, например, представляет собой почти монохроматический сигнал с периодическими биениями, и его частота точно совпадает с предельной скоростью распознавания меняющихся изображений, видимых человеком - кадровая развертка зрительного анализатора).

Итак, пусть есть сигнал, представляющий собой «осцилляторную смесь». Заметим, что если нам были известны колебательные характеристики осцилляторов, их количество и мощность, т.е. вклад каждого в результирующий сигнал, то этот сигнал можно будет повторить (моделировать) абсолютно точно. А если моделировать его с опережением по времени относительно основного генератора, то каждый отсчет можно будет точно предсказывать с любым желаемым упреждением (повторение истории с однорукими бандитами).

Если же характеристики осцилляторов неизвестны, то можно преобразовать сигнал в частотную область, а затем провести обратное преобразование с синтезом сигнала вперед, т.е. обеспечить предсказание во временной области.

В то же время возникает противоречие: «К несингулярным относятся процессы, для которых невозможно проследить характер причинно-следственных связей, так как они являются результатом суперпозиции большого числа элементарных процессов. Для них принципиально невозможно осуществлять прогнозирование мгновенных значений» [6]. Т.е. в соответствии с базовыми положениями современной теории невозможно собрать из большого количества элементарных монохроматических сигналов предсказуемый шум.

Возможно, конечно, что под элементарными процессами имеются в виду некоторые дельта (?) импульсы, не имеющие представления в частотной (NB!) области. Тогда возникает вопрос физической реализуемости этих компонентов, и призрак полосы пропускания, устремленной в бесконечность, тревожит воображение.

Вероятнее все же, что физически реализуемые, распространяющиеся по одному каналу связи, сколь угодно близкие к белому шуму процессы не могут быть такими, которые невозможно представить в частотной области. Для непредсказуемости этих процессов нужно будет иметь бесконечную полосу пропускания и соответственно бесконечную мощность источника сигнала. Если же полоса пропускания ограничена, а при дискретизации сигнала выполняется условие Котельникова (скорости Найквиста), то предсказание, хотя бы с большой погрешностью и очень малым упреждением, должно существовать (ниже будет показана иллюстрация ошибочности и этой гипотезы).

Теоретические соображения, а также собственные алгоритмы и программы полезно подтверждать экспериментом. Поэтому сначала синтезируем заведомо известные сигналы и проверим их спектр и функцию распределения.

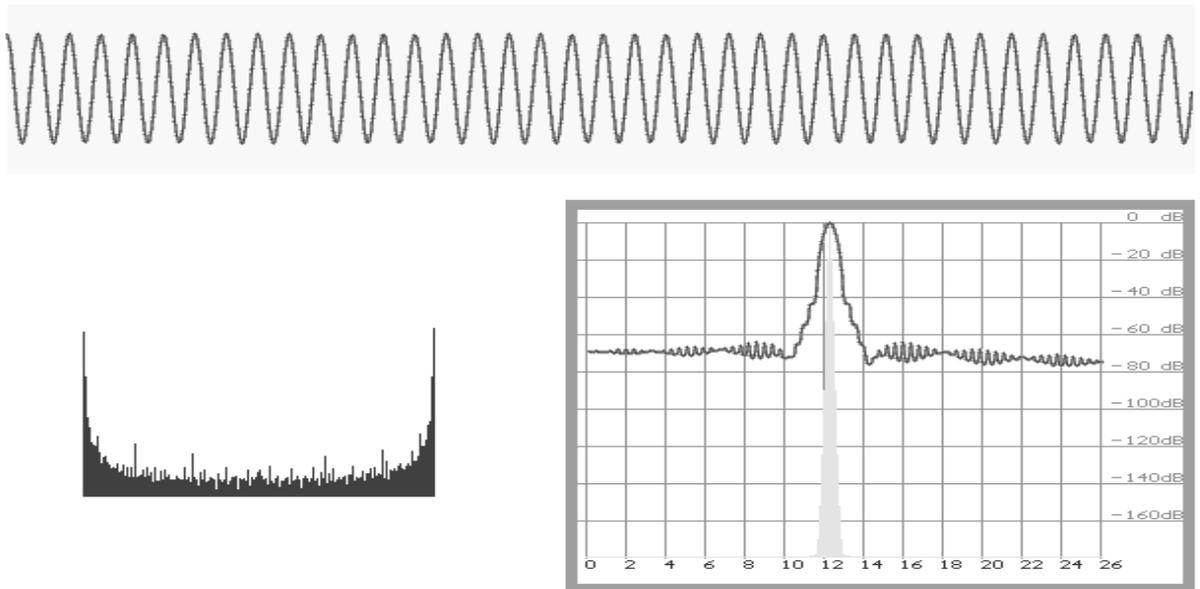


Рис.5. Монохроматический сигнал, его функция распределения и спектр (используется окно Хэмминга).

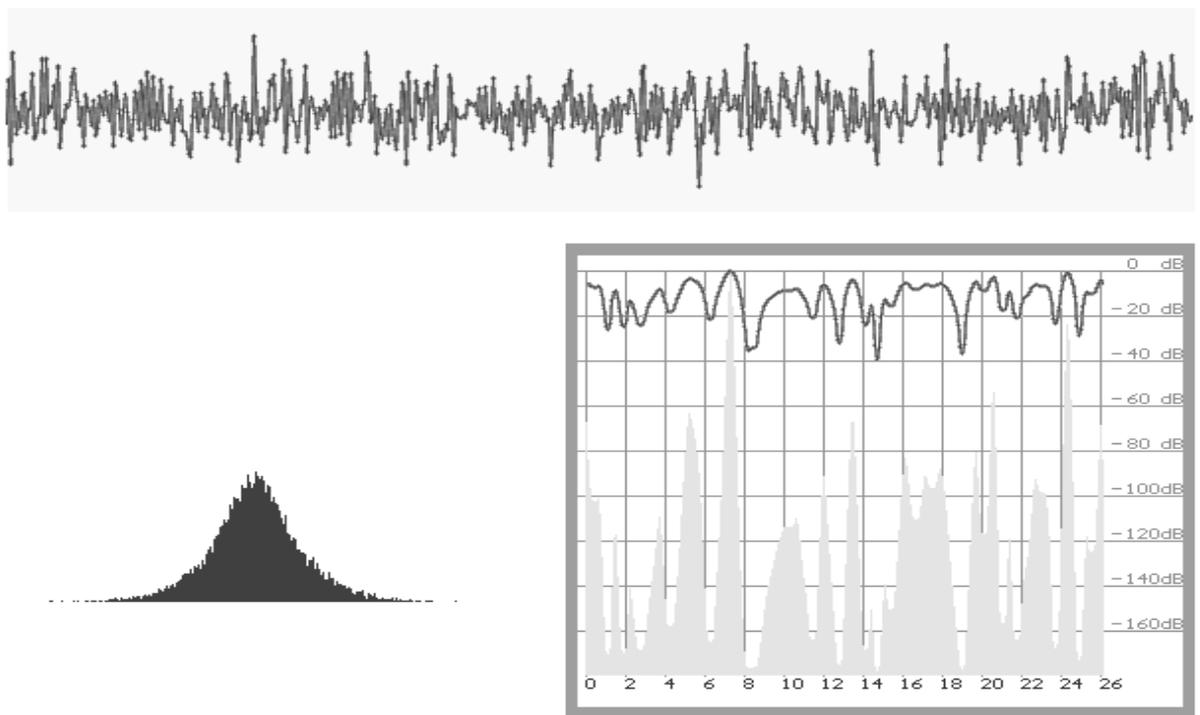


Рис.6. Нормально распределенный белый Гауссов шум (спектр вычислен для короткой реализации и поэтому видна неравномерность амплитудно-частотной характеристики выбранного фрагмента).

Предъявленные результаты моделирования вполне соответствуют базовой теории, т.е. можно ожидать, что алгоритм и программа работоспособны.

Теперь синтезируем сигнал, который хотел получить Кельвин – аддитивную смесь разнообразных генераторов монохроматических сигналов со случайной начальной фазой. Заметим, что взаимное влияние осцилляторов во временной области не будет иметь значения для результирующего спектра, т.к. полная энергия «системы из кабелей и маятников» будет оставаться почти постоянной при достаточно высокой добротности.

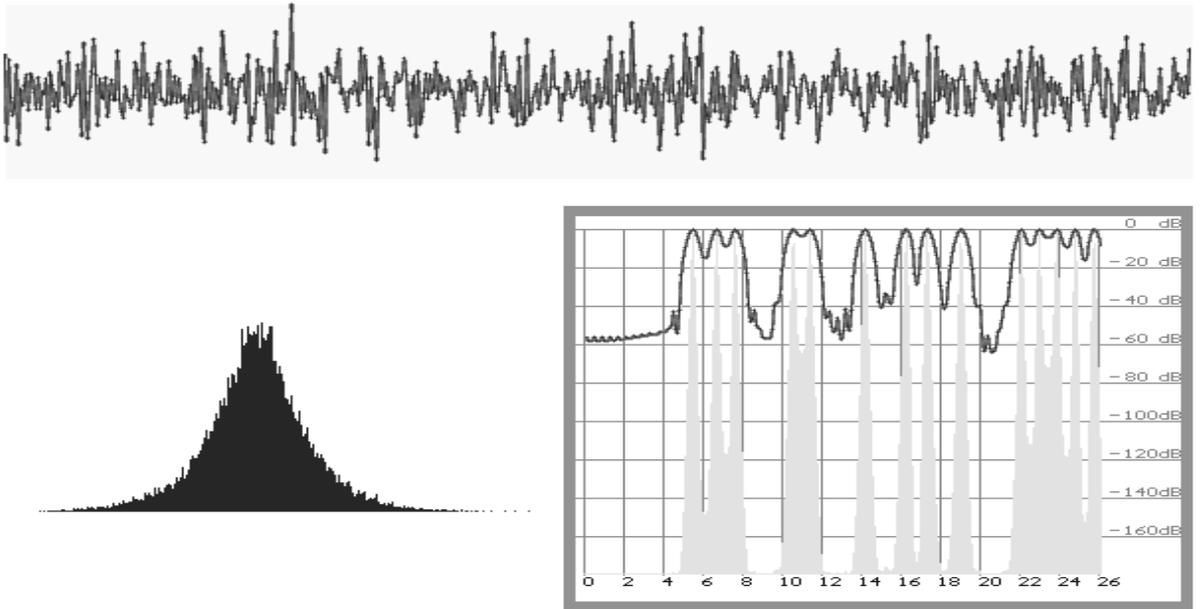


Рис.7. Шумовой сигнал, полученный в результате суммирования монохроматических спектральных компонентов с произвольной начальной фазой.

Можно ожидать, что этот сигнал будет полностью предсказуем, но следует также принять к сведению, что биения аддитивных монохроматических процессов обязательно приведут к периодичности результирующей последовательности. Период может стремиться к очень большим значениям, но он будет обязательно.

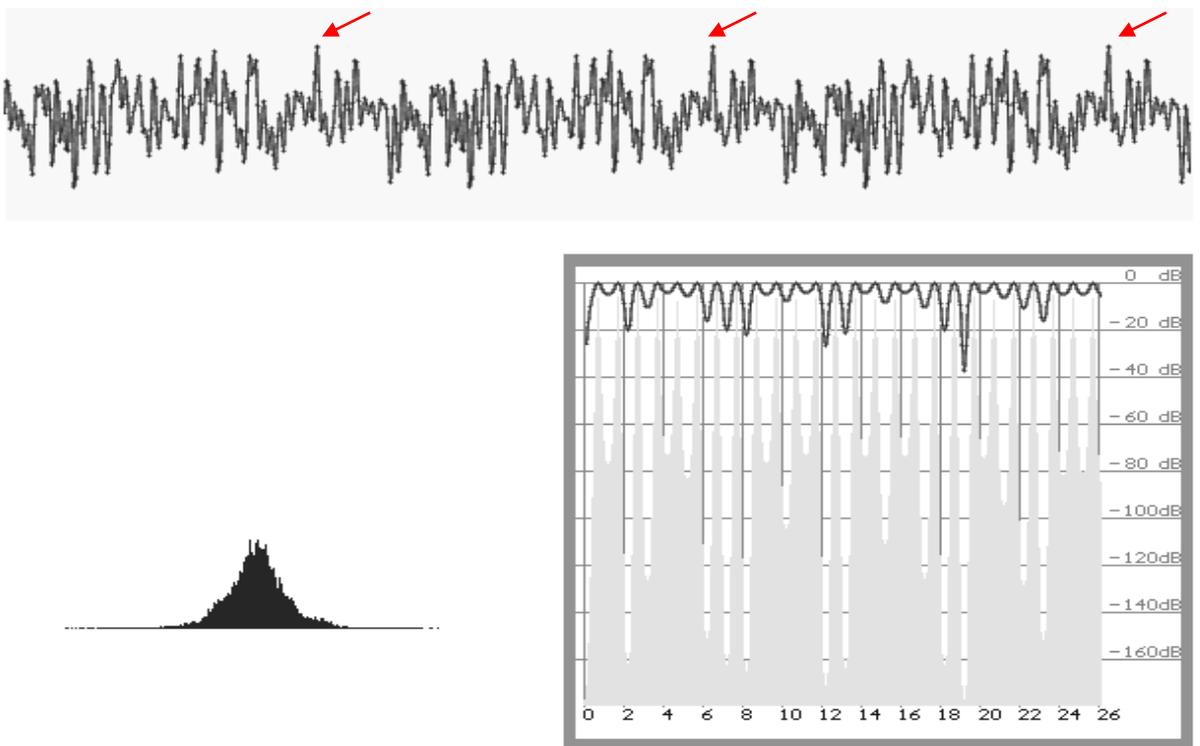


Рис.8. Иллюстрация предсказуемого шумового сигнала с равномерным спектром. Неудачные частоты резонаторов выбраны преднамеренно – во временной области получается эффект, похожий на красивую нестабильность Кельвина-Гельмгольца. Период повторения указан стрелками

Очевидно, алгоритмы синтеза предсказуемого шума и самого предсказателя настолько просты, что их может за полчаса запрограммировать любой рецензент научного журнала (не говоря о редакторе). Но для контроля все-таки проверим возможность эффективного предсказания монохроматического сигнала – здесь нет сомнений в теории, но проверка программного модуля никогда не бывает лишней.

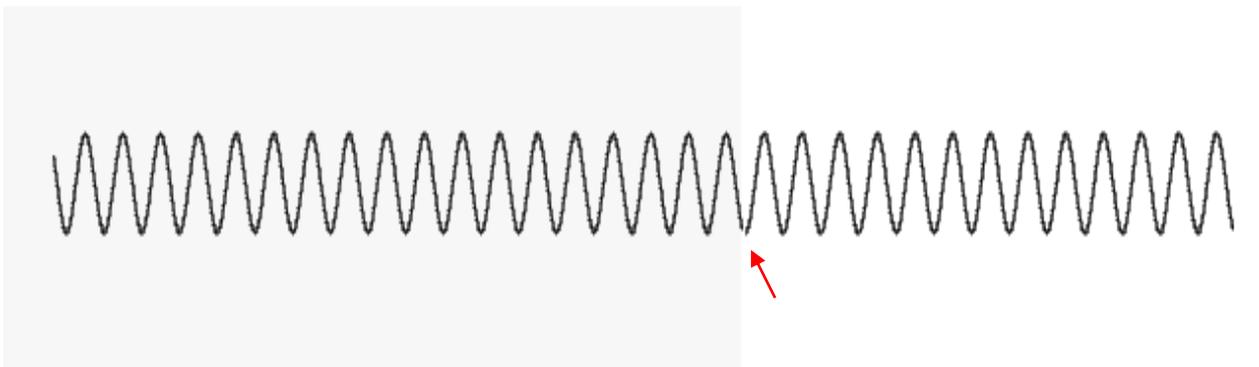


Рис. 9. Пусть есть гармонический сигнал, известный анализатору только в пределах левого поля, выделенного серым цветом. Решается задача продолжения сигнала при сохранении его амплитуды, частоты и фазы. Момент перехода к предсказанию показан стрелкой.

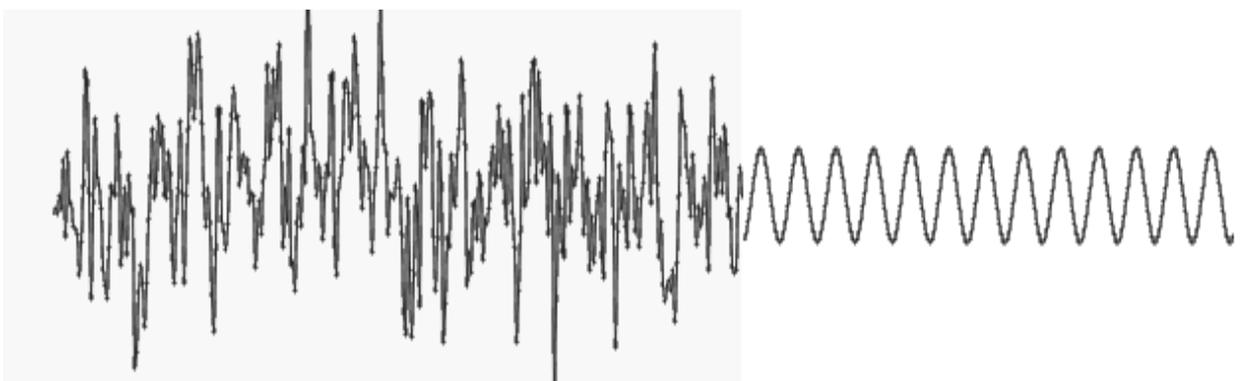


Рис.10. Смесь того же сигнала с гауссовым шумом. Обнаружение и выделение доминирующей на спектре одной монохроматической компоненты обеспечивается практически без искажений с очень небольшой частотной погрешностью.

После очевидно правильных иллюстраций последующая проверка должна включать точное (!) предсказание кельвиновского сигнала, иначе все предыдущие рассуждения стоят недорого, и их можно игнорировать.

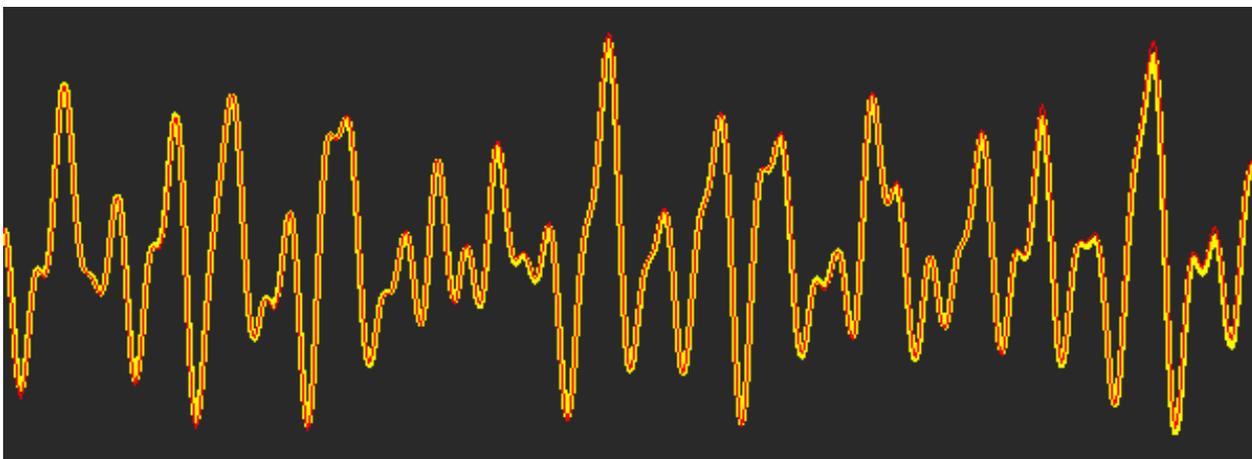


Рис.11. Реверсное предсказание шума Кельвина: желтой линией обозначен реальный сигнал, тонкой красной линией – предсказанные значения. В правой части графика видно небольшое ухудшение совпадения (вследствие погрешностей окна данных спектроанализатора и неполного подавления боковых лепестков). Скорость развертки и коэффициент усиления увеличены относительно предыдущих иллюстраций для наглядности.

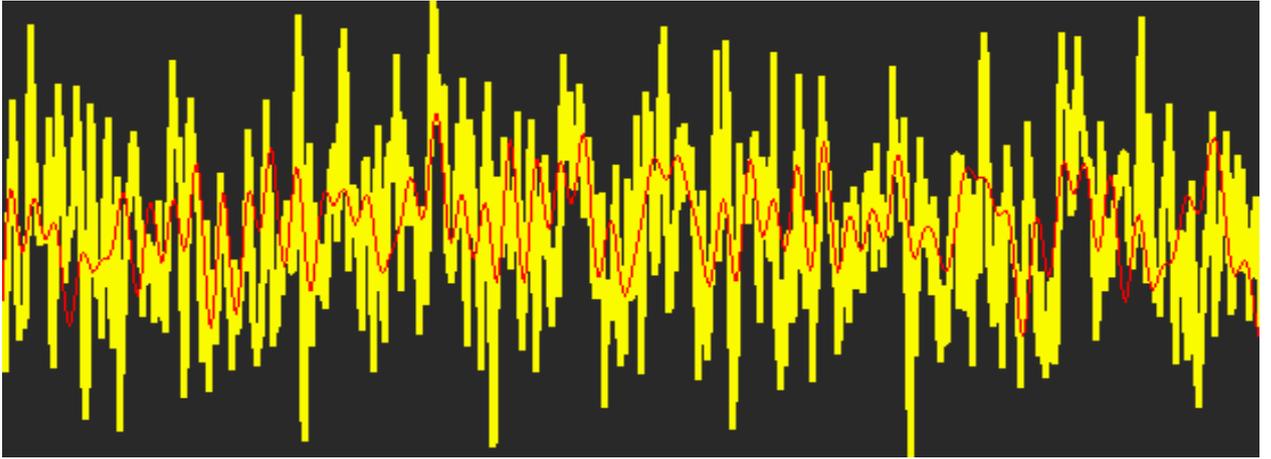


Рис.12. Реверсное предсказание ограниченного в полосе шума Гаусса. Явная невозможность решения задачи: в какой-то мере отслеживаются только низкочастотные компоненты колебаний. Высокочастотная часть сигнала уходит в режим «маскировки спектров» также вследствие нарушения условия Котельникова.

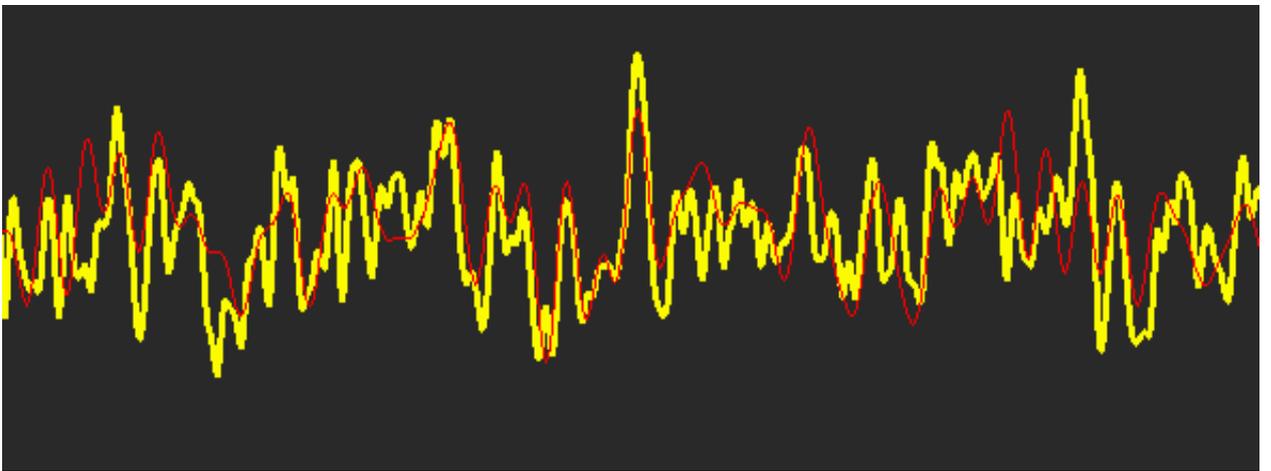


Рис.13. Дополнительное ограничение полосы шума Гаусса сверху. Предсказание по-прежнему проблематично, хотя рассогласование входного и предсказанного сигналов заметно уменьшилось.

По результатам проведенных экспериментов можно сделать два предварительных вывода.

1. Кельвиновский шумовой процесс хорошо предсказуем, по меньшей мере, на короткой реализации.
2. Процесс Гаусса не удастся эффективно предсказывать даже при условии ограничения полосы пропускания сверху.

Естественно, возникает вопрос: зачем нужны эти упражнения с предсказанием, если существование шума Кельвина до настоящего времени не является общепринятым и, возможно, иллюстрирует только навязчивую идею автора (чего никогда не следует исключать априори). В оправдание приведу цитату: *«Отнесение процесса к классу случайных может быть обусловлено либо его физической природой, либо условиями его изучения, приводящими к недостаточности априорных данных [6]».*

И вспомним, например, что в нашей среде обитания работает множество разнообразных двигателей внутреннего сгорания. Практически все они имеют на выходе вращательное движение со стабильной или относительно медленно меняющейся частотой. Колебания, возникающие под влиянием этого движения, имеют множество четных и нечетных гармоник, обусловленных как вибрациями вследствие некомпенсированных моментов инерции, так и работой клапанного механизма и приводов отбора мощности. Кроме гармоник генерируются также рассинхронизированные шумы.

Результирующий акустический сигнал имеет как предсказуемый компонент, состоящий из шумов и вибраций, синхронизированных с основной гармоникой, так и непредсказуемую «чисто гауссовскую» часть. Это может дать возможность оценки

технического состояния двигателя, т.к. записанную реализацию шума двигателя можно предсказать в инверсном времени.

В результате получится две реализации сигнала: изначальный шум и результат его предсказания. Вычисляя коэффициент $R(p)$ (нормированной взаимной корреляции) между этими двумя последовательностями, можно оценить вклад кельвиновского компонента в суммарный шум.

И если, например, коэффициент корреляции растет в процессе эксплуатации, то это несомненный признак износа или повреждения деталей, перемещающихся синхронно с основными гармониками. Если же коэффициент корреляции уменьшается (растет непредсказуемый компонент), то где-то есть повышенное трение, пропуск газов и т.п. Особенно полезным такой метод может быть для турбин непрерывного цикла – постоянный бесконтактный контроль лишним не бывает.

Разумеется, вместо нормированной корреляции можно вычислять энергию рассогласования, или вообще воспользоваться всем набором инструментов математической статистики. Но как мне кажется, получить много дополнительной информации относительно корреляционного анализа вряд ли получится.

Допустим, что параметр, соответствующий значению $R(p)$, т.е. «коэффициенту предсказуемости сигнала», может найти применение в самых разных областях. Даже биржевому игроку или валютному спекулянту интересно будет иметь прогноз на вероятность выигрыша по предыдущему поведению кривой цен. О таких областях, как электроэнцефалография, можно и не упоминать: получение карты аномалий совокупности мозговых ритмов, несомненно, повысит уровень диагностики. Особенно в тех случаях, когда ЭЭГ изменяется неявно, так, что только очень опытные специалисты позволяют себе предполагать патологию.

Вероятно, оценка предсказуемости сигналов также может быть использована и для исследований в радиоастрономии, которая изучает как чисто стохастические сигналы, так и ритмичные послылки разнообразных источников. Да и на «карту предсказуемости неба в радиодиапазоне» стоило бы посмотреть.

Кроме того, хорошо бы узнать: является ли гравитационный шум процессом Кельвина, и отличаются ли от него мощные гравитационные волны?

Для этого примем как аксиому фундаментальный запрет: «Если сигналы существуют одновременно и их спектры перекрываются, то полное разделение сигналов невозможно». Учитываем также, что зарегистрированная детекторами гравитационная волна всегда существует одновременно с гравитационным (и не только) шумом. Разумеется, разделить эти два сигнала полностью не получится; спектральное отбеливание просят не предлагать по очевидной причине: наличие априорной гипотезы. Естественно, не подходит и классический корреляционный метод, который чувствителен к фазе анализируемых процессов. Но если вычислять коэффициент предсказуемости сигнала, то описанные препятствия можно обойти, не нарушая базовых положений теории информации. Попробуем поступить именно так, и начнем с контроля алгоритма фильтрации.

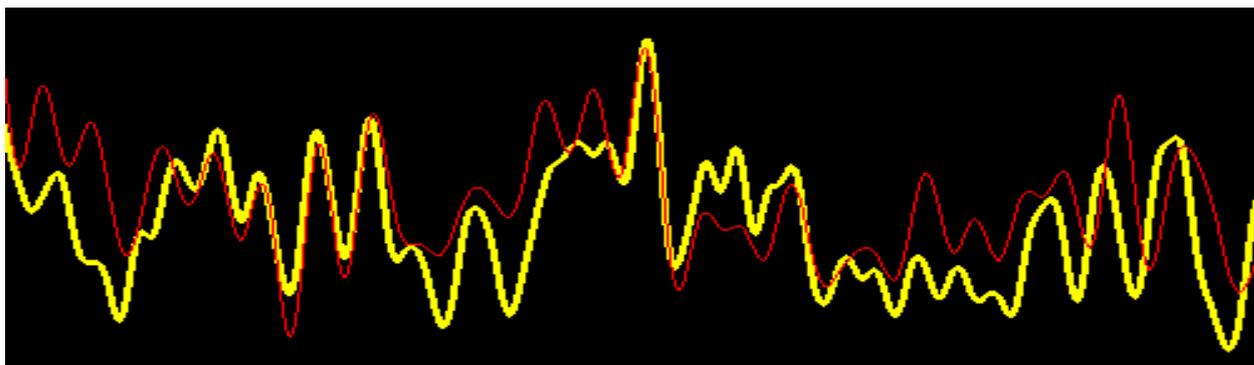


Рис.14. Контрольный сигнал шума Гаусса после прохождения системы фильтров, идентичной той, которая применялась для выделения гравитационной волны. Реверсное предсказание. $R(p) = 0.5035$

Используем записи проекта LIGO и не забудем, что выделение полезного сигнала из самых разнообразных помех происходило на пределе возможностей аппаратуры. Естественно, «зашумленный» сигнал должен будет показать меньшую предсказуемость.

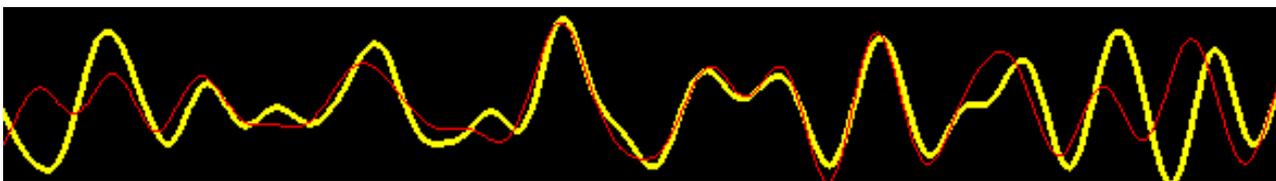


Рис.15. «Спокойный» беспомеховый участок записи гравитационного фона. $R(p) = 0.6608$

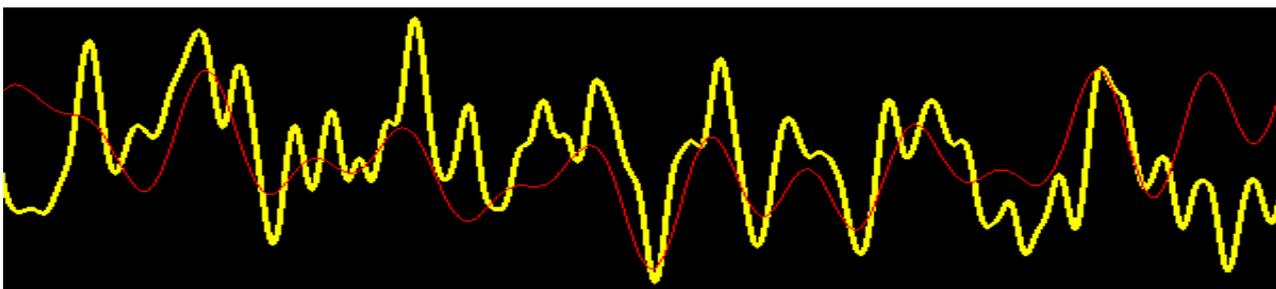


Рис. 16. Участок записи с какими-то внешними помехами. $R(p) = 0.3162$

Результат вполне ожидаемый: чем больше шумов и помех, тем ниже коэффициент $R(p)$. Теперь попробуем предсказать официально подтвержденную гравитационную волну. Еще раз отметим, что ее выделение [7] произведено из любезно предоставленных для общего пользования нефильтрованных записей LIGO.

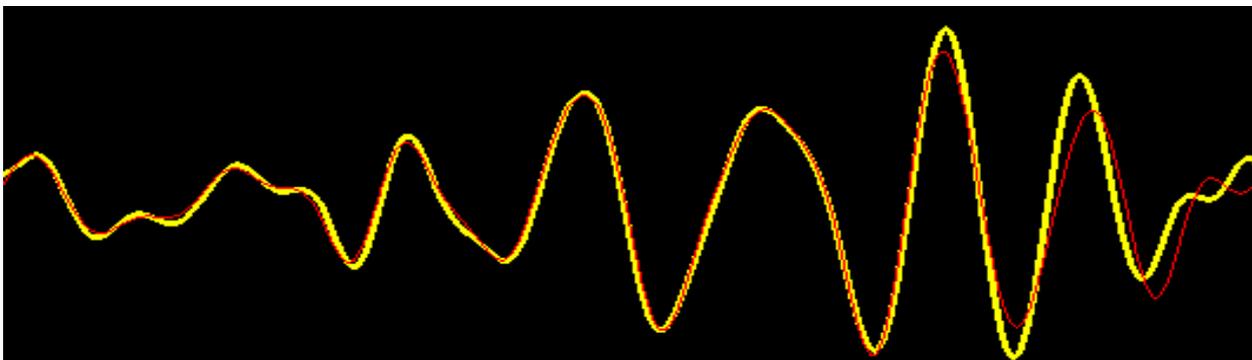


Рис.17. Реверсное предсказание гравитационной волны, выделенной из записи лабораторий LIGO с помощью согласованной КИХ фильтрации без использования (NB!) метода спектрального отбеливания. Коэффициент предсказуемости сигнала $R(p) = 0.9587 (!!)$

Полученное значение коэффициента ожидаемо велико: ведь исходным объектом была система масс, вращающихся с увеличением угловой скорости. И это вращение обеспечивает главный косинусоидальный компонент в сигнале, зарегистрированном на расстояниях много больших радиуса вращения.

Неудивительна также высокая предсказуемость гравитационного фона: действительно, вращающихся масс в зоне пространства, откуда может приходиться достаточно мощный сигнал, очень много. И эти осцилляторы должны формировать именно шум Кельвина – аддитивная смесь сигналов вращающихся масс вполне соответствует его системе из «...маховиков, брусков, пружин». Впрочем, подобные гипотезы в области гравитационной астрономии отнюдь не приветствуются состоявшимися специалистами в этой области.

Главное, что $R(p)$ действительно работает, и результаты вполне соответствуют ожиданиям. Его можно использовать на практике. Игроков на бирже что ли порадовать?

P.S. Для вычислений $R(p)$ применен алгоритм Пирсона. Он имеет одну неприятную особенность. Если в подкорневом выражении появляется «0», то возможна ситуация деления на него с соответствующими последствиями. Устранить этот недостаток можно принудительным добавлением к подкорневому выражению одной единицы младшего разряда, точность вычислений практически не страдает.

Список использованной литературы:

1. http://loveread.me/view_global.php?id=76391
2. <http://tredex-company.com/ru/krasnoe-smeschenie-kak-universalnyj-priznak-porazheniya-serdechno-sosudistoj-sistemy->
3. Подтверждение тезиса, который объяснял своим студентам хирург Евгений Юрьевич Крамаренко: «Больной выздоравливает благодаря лечению, несмотря на него и вопреки лечению».
4. Бендаг Дж., Пирсол А. Прикладной анализ случайных данных. М., Мир, 540 с., 1989
5. <https://cyberleninka.ru/article/n/modelirovanie-neustoychivosti-kelvina-gelmgoltsa-modifitsirovannym-metodom-diskretnyh-vihrey>
6. vestnik.mauniver.ru/v09_3_n23/articles/18_prokh.pdf
7. <https://arxiv.org/abs/2212.05851>