CENTER FOR COMPUTER RESEARCH IN MUSIC AND ACOUSTICS
MAY 1988

Department of Music
Report No. STAN-M-48

SIMULATING PERFORMANCE ON A BOWED INSTRUMENT

Chris Chafe

CCRMA
DEPARTMENT OF MUSIC
Stanford University
Stanford, California 94305

Simulating Performance on a Bowed Instrument

Chris Chafe

Center for Computer Research in Music and Acoustics (CCRMA)
Department of Music, Stanford University
Stanford, California 94305

Introduction

The accompanying sound examples demonstrate music synthesized from a physical model and controlled by simulated physical performance gestures. This paper presents some issues in designing a control system for such a synthesizer. The system that has been developed runs considerably slower than real-time and is intended as an environment for answering some questions about live control of an eventual real-time system. The control methods in a playable real-time instrument are similar to those used here, in which the computer interprets the musical score.

The attempt to simulate performance on bowed instruments actually stems from an interest in developing new instruments, rather than pure simulation. A good simulation serves as a useful "benchmark" that provides a starting point for exploration of novel instruments and musical materials for compostion.

The control method is specific to synthesis from a physical model of a bowed string. Five time-varying control signals are generated for the model: string length, string damping and bow speed, pressure and position. Other instrument types could be tested using a similar approach. Articulation rules would be adapted to the particular parameter sets of their synthesis models.

Time-varying envelopes for the parameters are created by concatenation of short envelope segments corresponding to performance gestures. Gestures themselves are not represented in actual physical terms, so there is no notion of the extent or rate of motion of the fingers or bow. Instead, the system represents gestures in terms of their effect on the string parameters.

Musical scores are coded as lists of gestures. The method is a variant of tablature notation, in which a score is described from the point of view of hands manipulating the instrument, rather than pitches on a stave. For example, events described in this fashion include *martele attack on string III* in the right hand or *hammered pitch at the fifth on string I* in the left hand.

When creating a performance, the system adds details that cause gestures to conform to the changing state of the musical phrase. The intended result is consistency within complex gestures that perform multiple notes as well as some amount of expressiveness.

Synthesis Method: Physical Modelling

At the outset, it should be mentioned that though computer music synthesis using physical models is promising, it is still largely a theoretical field. Commonly restricted to running in software, sound generation is many times slower than real-time *. Real time systems require either new types of synthesizer hardware (which are likely to be built from standard digital signal processing VLSI) or general purpose super computers. The chapter by Wawrzynek addresses one physical modelling approach and specially-designed hardware that implements the models in real time [Wawrzynek, 1988].

---

* To date, no composition that I'm aware of has been made using the technique.

The bowed string algorithm used in this paper runs in software. Since the purpose here is to discuss control of the cello model, background concerning the algorithm itself should be found by referring to the following sources: [McIntyre and Woodhouse, 1978, McIntyre, et al. 1983, Smith, 1983, Cremer, 1984, Weinreich, 1985, Smith, 1986].

A physical model of an acoustic source can represent a vibrating system of arbitrary complexity. A complete cello-like synthesizer is constructed by coupling together a number of bowed string models that simulate each string's internal reverberation and the bow's frictional driving mechanism, and which are further coupled to the bridge, the sound box and the air. Weinreich has termed the bowed string algorithm a method for "synthesis from first principles." It is a simplified, but nonetheless accurate, description of a dynamic physical system. The algorithms representing the four cello strings are iterated in time from some initial state, eg. an open string at rest. As the bow excites string motion, waves begin to circulate that are emitted as sound at the model's output. Events in the musical score, through the intermediary control system, cause changes to parameters controlling the bow and string.

A cellist develops skill at manipulation of five basic parameters. *String length,* controlled by fingers on the player's left hand, determines the round-trip time of the recirculating waves and the resulting pitch of the sound. The right hand controls *bow speed, pressure and contact position,* affecting loudness and tone quality. A fifth control, *string damping,* controls the amount of wave recirculation and varies with movement of the fingers and bow touching the string.

In synthesis experiments with a physical model of the cello, a number of advantages have become clear:

- **Regimes of oscillation are correct in the time domain.** Self-sustained oscillators such as bowed strings are *dynamical systems.* The system's *attractor* (equilibrium state) under normal bowing conditions is a waveform exhibiting *Helmholtz motion,* named after the acoustician who first described the periodic mechanism of string capture and release by the bow. Figure 1 compares real and synthesized bowed waveforms.
- **As in the real instrument, transient behavior of the system is state-dependent.** Response to a particular articulation depends on what the system was doing in the recent past. Sound example number 63 demonstrates multiple stikes on a physical model of a bell. The repeated strikes sound different because of their interaction with the vibrating system.
- **The model has the same external controls as the physical system.**These are the parameters a cellist develops skill at manipulating.
- **There is an "intuitive feel" to the system's response.** A cellist can recognize and imitate a synthetic articulation and recommend improvements in the control values. The complete range of cello articulations can be synthesized.

The most important restriction concerning control of the model, is that the model's state should persist through successive articulations and events, allowing the system to produce real-sounding transients. Traditional speech or music synthesis techniques, eg. linear predictive coding, formant synthesis, additive synthesis, frequency modulation or sampling methods, don't easily produce real-sounding transients. In physical models, finely detailed transients are achieved when new events interact with the reverberation of previous events. A particular articulation, played twice, may sound different as it interferes with a system that is already in motion. The state of the system at any given point in time is a result of complex interactions between external control parameters and recent system "memory." Due to this accurate transient behavior, dynamical system models are useful in a wide range of real-world simulations [Campbell, 1985].

This accuracy creates, in music synthesis, recognizable instruments. A wide range of acoustic cues contribute to the dynamic signature of an instrument in play, especially features that are co-varying. It doesn't matter particularly *what* is being performed. For instance, the model can produce tones or passages that are recognizably the sort produced on the instrument by unskilled players. Refining the performance of the control system is reminiscent of early exercises on the instrument. When something is amiss, the best approach has been to compensate the controls as if listening to the actual instrument. Tone quality, for example, might be improved by specifying something like, "...use more bow pressure at that point."

Scores: Segmenting Performance

Music is performed on the cello synthesizer by a control system that replaces the cellist. Its rules are gesture-based, and imitate the effect of the player's actions on the strings. The system also attempts to reproduce some of the interpretive functions of the player such as musical phrasing. Time-varying envelopes for the synthesis parameters are calculated by rules that correspond to basic gesture segments. Usually, a few of these are required to form the complete envelope of a musical note.

The system operates from musical scores coded in common notation and enhanced with explicit markings for bowing and fingering (to choose which string a particular note is played on). These markings are often added by players to their parts, indicating fingering and bowing choices for performance of a particular passage. Each succeeding pitch is associated with either a fingering change, bow change, or both. The old forms of tablature style scores for lute are reminiscent of this approach, in that they notated the placement of the hands on the strings. Figure 2 illustrates an example coded as input for the system.

The gestures that the hands can perform vary in complexity. The simplest items are those that initiate a bow stroke or finger a new pitch. More intricate operations are possible which result in coordinated activity across multiple strings, such as rolled, or "broken," chords. The system breaks each gesture into strings of smooth sequences of short envelope segments, for example, those resulting in string release, bow accleration, pitch sustain or vibrato. The larger gestures described by the musical notation, such as a note sequence, trill or slur, are formed by compounds of these "atomic" segments.

For each gesture in the score, sequences of segments are cued up in a time-ordered list by the control system. In most cases this sequence contains one or more transition segments paired with a sustain. For example, a change of bow is comprised of a quick deacceleration, a re-attack and a sustain segment.

The extendable duration of sustain segments distinguishes them from transitions, whose durations are fixed. Sustainable segments prolong the string state, possibly modulating it vibrato or tremolo. The intervening transitions determine trajectories for the hands as they move between sustainable motions. For example, left hand transitionals include sliding, hammering, or homing in on a new pitch.

Each segment contributes to the envelopes of a few synthesis parameters. The number of parameters that will be inflected synchoronously depend on the particular segment. In a reversal of bow direction, the re-attack segment affects four parameters: bow speed, pressure, pitch and damping.

Several kinds of phrasing marks are likely to be encountered in a score and are referred to as *phrase controls.* These include dynamic level, vibrato rate, tone quality, and tempo, among others. Separate slow changing envelopes which are internal to the system are generated from marks in the score. Their levels represent the *phrase state* at any given instant.

Phrase controls are updated by an ever-present background process. These values are relatively stable since they change over the course of seconds or tens of seconds. Though a given control level usually persists longer than a single note, exceptions can occur, for instance, as in a *messa di voce,* (a prolonged tone with a swell in dynamic level).

The control system can insert unscored segments when interpreting a score. Some may belong to a particular playing style while others arise from a need to simulate a natural level of "sloppiness." Citing some inaccuracies common in string playing: pitches are rarely placed directly on target by the left hand and spurious sounds often accompany position shifts or string crossings. Computer-perfect synthesis without a dose of nature's noises has a lifeless quality.

Simultaneity: Multiple Articulators

Novice players are familiar with the burden of the many controls that they must track: the bow, intonation, tempo, loudness, and timbre. In effect, these items can be as components of a "polyphonic" texture, containing several independent parts, or "voices," for which the hands are the articulators.

From the instrument's point of view, the two articulators that act on it are sometimes coordinated and sometimes completely independent. Both hands simultaneously contribute to the value of a synthesis parameter. This is clearly the case for the damping coefficient in the string, as well as for string length and the bow parameters.

Two independent processes represent the hands in the control system. These processes will be assigned different kinds of gestures that invoke string excitation, and pitch or damping change. Regarding string excitation, besides the right hand bowing the string, the left hand can hammer down a new pitch and either can pluck. Given the possibility for left hand sounds, the current system allows the left hand process to add excitation by inflecting the bow parameters. A hammered pitch is simulated by a sharp jerk in the bow envelopes.

Updates to the synthesis parameters occur every 10 milliseconds. Prior to each update, the hand processes evaluate their contribution to the control envelopes and their respective contributions are summed for each of the synthesis parameters.

Segments in the two hands may be changed synchronously or asynchronously depending on the type of musical material. In typical performance, this texture may switch rapidly from a one-to-one correspondence, to many-to-one or entirely independent relationships between the hands. For example, in Figure 3, several segments have been generated by the system. Due to the initial three-note slur there are more segments for the left hand than the right.

An additional parallel-executing process evaluates the state of the phrase controls. The hand processes evaluate their segments within this context. The desired effect of this mechanism is that a given gesture or articulation will translate into different envelope shapes depending on the dynamic level or tone quality.

Control processes, synthesis processes, gestures and segments are objects in the system's computer program. Each has an associated set of execution rules. At initialization time, the instrument configuration is declared giving the number of strings and their tunings. The system then copies the required number of each object prototype. For the sound examples, from 1 to 4 strings were played by 2 hands (of course, being a computer, additional hands can be made available for more difficult scores).

All processes execute each tick in an order that follows a hierarchical pattern *. The instrument is the *father* process and hand, phrase and string synthesis processes are *offspring.* The father contains a scheduler that cues the time-ordered segments into its offspring processes. The evaluation order (at each update) is as follows: instrument, phrase controls, and then string by string: left hand control and right hand control. The updated value is sent to the string synthesizer process which then outputs 10 milliseconds of sound.

Rule-based envelope generation

Envelopes are calculated by the segments' rules rather than fetched from lookup of envelope tables. There are two principal reasons for taking this approach:

- Multiple synthesis parameters can be inflected in a coordinated way.
- The rules automatically adjust envelopes according to the changing phrase state.

Recorded measurements of violin bowing gestures support the notion that multiple envelopes are often inflected by one articulation [Askenfelt, 1983, 1986]. Performance of a bow stroke with a certain desired loudness and tone quality requires balancing of the three bowing parameters. Bow speed, pressure and contact position are interdependent in a way that is best understood graphically [Schelleng, 1973]. Plotting these quanitities in separate dimensions, there is a locus of "normal tone" in which these quantities are

---

* (a structure inherited from the original implementation in the FORMES environment for musical control tasks [Rodet, 1984]).

balanced for different dynamic levels. Outside of this region the bowed effects of sul tasto, sul ponticello, flautando and even undertone (octave splitting) are found. The desired balance of the bowing parameters is determined by rules contained in the bowing segments, as demonstrated in the third sound example. It demonstrates variation of dynamic level and contact position while repeating the same articulation.

The rule-based method for envelope generation facilitates coordinating shape and timing of the different articulations. For example, all five synthesis parameters are perturbed together by a *flying spiccato* bow stroke. Bow speed, pressure and contact position as well as pitch and damping are affected by the abrupt attack as the bow drops on the string.

The following as an illustration of the detail that can be embedded in segment rules: as mentioned before, the left hand excites a small amount of string motion in hammering down on a new pitch when striking the fingerboard hard enough. In the *hammmered pitch* segment, a short interval of heavier damping occurs during which the finger stops the string just before it's fully seated on the fingerboard. This brief muting is followed by new excitation from the hammered strike. Similarly, *pulling off* to a new pitch is executed by a segment that includes a left hand pizzicato.

The use of a "global" phrase state also provides a means for regulating details of envelopes affecting individual strings. Multiple-string gestures such as bowed triple-stop chords can be regulated by rules that lookup phrase attributes. In this example, rolled chords are played across three strings. Since the bow is limited to touching two strings at a time, chords are played as a succession of diads. Bowing envelopes that are calculated for the individual strings track the current dynamic level and tone quality to behave in a mutually consistent fashion. Figure 4 depicts the energy through time on each of the 4 strings as a country fiddle tune is played (sound example 2).

Certain articulations create strong multi-parameter accents. A *sforzando* attack would generally affect vibrato as well as the bow. Such accents are engineered as brief boosts to phrase controls, triggered by the accent segment. Transient "bumps" are imposed on envelopes for timbre, dynamic and vibrato are short lived and decay exponentially.

Bow Articulations

Segments available in the rule-base of the system include representatives from a small "taxonomy" of bowed articulations. Five features have been taken into account in classifying them:

- **Reversal of Bow Direction.**This distinguishes *detaché* from *legato* bow strokes.
- **Initial displacement.**Significant bow pressure before the bow starts to move causes a hard attack. The initial string snap that results is sometimes called a *quasi-pizzicato.* Examples are *martelé, marcato and staccato.*
- **Pulsed accent.**Soft attacks have no discernable displacement, but can have some degree of accentuation after the string is in motion, as in *lancé, porté and louré* bowing.
- **Duration of separation.**Tones separated by small amounts of silence include *martelé, spiccato and staccato.*
- **Off-string bowing.**Dropped, lifted or ricocheted strokes are distinct from completely on-string tones. During the decay segment, off-string tones ring where their on-string counterparts are damped. *Piqué, spiccato, jeté and flying staccato* strokes are played off-string and are distinuished by amount of contact at attack and decay.

The fourth sound example compares several articulations whose envelopes are plotted in Figure 5.

Summary

The control system has been tested in several different musical situations, both with hand entered scores and scores composed by algorithmic means. In experimenting with different instruments, envelope-generating rules have been created that simulate guitar and koto performance. The rule base can be expanded for a selection of musical styles, instruments or even the personal touch of different performers.

The rules currently in the system have been refined through repeated synthesis trials, in the absence of measurement from actual playing. In the future, such measurement is possible through instruments which can sense bowing and pitch information. An example is the system of Askenfeldt's. Under development at CCRMA are string family MIDI controllers which will allow analysis of live performance, and which will help to determine more precise envelope-generating rules.

Note lists, which are the most common data structure for specifying computer music scores, would be inappropriate for synthesis from physical models of continuous sounding instruments like the cello. Control of the instrument involves several parallel processes whose synchrony varies depending on the musical texture. Slurs are a simple example of asynchrony between the two hands of a cellist. Subsequent notes continue the sound before the previous note is extinguished. In the note list style of control, only separate and isolated notes can be sounded and slurs are approximated by keyboard style note overlap. Moreover, the state of the synthesizer is re-initialized each note, eliminating the advantages of physical modelling for synthesis of natural-sounding transients.

Recapping the issues that have been considered in the design of the cello control system: Left hand and right hand gestures are merged in a way that permits their effects to interact in the synthesis. Gesture segments can coordinate changes in several parameters at once and track evolving phrase conditions. The synthesizer's state is preserved allowing the physical model to be put to full advantage.

The author wishes to thank Julius Smith, Xavier Rodet and Gabriel Weinreich for contributions of advice and software. The system was first implemented in 1984 at IRCAM Paris, France and again in 1987 at CCRMA.

Simulating Performance on a Bowed Instrument

Sound Examples, Chafe.

The first 4 synthesis examples were completed in 1988 using INTERLISP-D software with array processor support. For simplicity, several limitations were imposed: only a single bow point was computed (as though the bow had one hair), bow changes did not invert waveform phase and no body transfer function was used (just the raw string vibration was output, slightly reverberated). The last example, completed at IRCAM in 1984 was computed with an FPS-100 array processor under control of the FORMES version of the software.

- **Tuning Up**Simulating a cello soloist, this example is a good test for the entire system. Any "off-by-one" programming error and it won't tune up.
- **Fiddle Tune**A country fiddle tune is played in the cello register. The score was taken from a transcription of the tune "Wrassled with a Wildcat" which appears in "The Fiddle Book," in scordatura (with strings tuned A-E-A-E).
- **Repeated Tones with Dynamic and Position Changes** A diad is played on two strings with changing dynamic level and contact position of the bow.
- **Articulations**Several synthesized bowed articulations are compared:

  1) *"NOVICE"* A squeaky stroke resulting from an imbalance of bow speed vs. pressure.

  2) *MARCATO* An accented bow change.

  3) *LOURÉ* A lightly accented *legato.*

  4) *MARQUÉ* A heavily accented *legato.*

  5) *MARTELÉ* A separated, accented stroke (*detaché).*

  6) *SPICCATO* A lifted bow stroke.

  7) *LANCÉ* A softer *spiccato* .

  8) *PORTÉ,* Soft attack, with a delayed accent.

  9) *"IMPOSSIBLE PORTÉ" Porté* with an abrupt mid-stroke accent.

- **Automatically Improvised Hot Jazz**The player system was hooked up to the output of a melody improvisation algorithm. Using a set of rules designed by André Hodeir, the algorithm created a violin line to fit a set of chord changes that were given as input. The guitar is accompaniment uses the same synthesis model as the violin.

References

A. Askenfelt, "A Simple Device for the Simultaneous Registration of Bow Motion and Bow Force," *Proceedings of the Stockholm Music Acoustics Conference,* 1983.

A. Askenfelt, "Measurement of Bow Motion and Bow Force in Violin Playing," *J. Acoust. Soc. Amer.*, vol. 80, no. 4, pp. 1007–1015, Oct. 1986.

L. Cremer, *The Physics of the Violin,* MIT Press, 1984.

D. Campbell, et al. "The Role of Computation in Nonlinear Science," *Communications of the ACM,* vol. 28, no. 4, pp. 374–384, 1985.

M. E. McIntyre, R. T. Schumacher, and J. Woodhouse, "On the Oscillations of Musical Instruments," *J. Acoust. Soc. Amer.*, vol. 74, no. 5, pp. 1325–1345, Nov. 1983.

M. E. McIntyre and J. Woodhouse, "The Acoustics of Stringed Musical Instruments," *Interdisciplinary Sci. Rev.,* vol. 3, no. 2, pp. 157–73, June 1978.

X. Rodet, "FORMES: Composition and the Scheduling of Processes.," *Computer Music Journal,* vol. 8, no. 3, pp. 32–50, June 1984.

J. C. Schelleng, "The Bowed String and the Player," *J. Acoust. Soc. Amer.*, vol. 53, no. 1, pp. 26–41, Jan. 1973.

J.O. Smith, "Techniques for Digital Filter Design and System Identification with Application to the Violin," *Department of Music Technical Report STAN–M–14.,* Ph.D. Thesis, Department of Electrical Engineering, Stanford University, Stanford California, June 1983.

G. Weinreich, "Violin Sound Synthesis From First Principles," *abstract of paper presented to the Acoustical Society of America Conference, J. Acoust. Soc. Amer.*, vol. 74, pp. sup. 1S52, 1983.

G. Weinreich, "Sound Hole Sum Rule and the Dipole Moment of the Violin," *J. Acoust. Soc. Amer.*, vol. 77, no. 2, pp. 710–18, 1985.

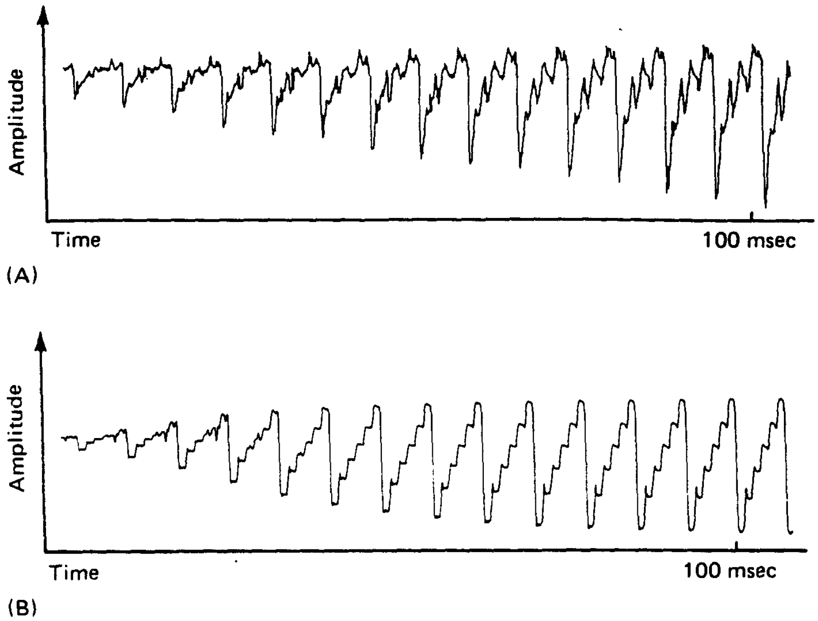
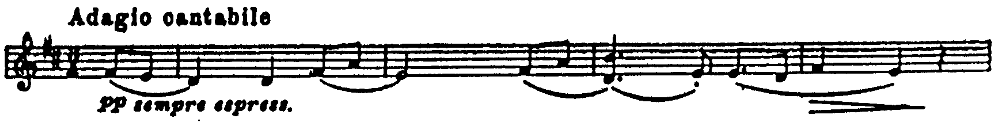


Figure 1
(A) sampled waveform and (B) synthesized waveform. The plots compare sampled and synthesized bowed cello waveforms. The upper plot was recorded off an isolated string in the *Celletto*, a bodyless electric cello with sensors implanted in the bridge directly beneath each string. Both traces exhibit sawtoothlike *Helmholtz motion* at the fundamental period modulated by crumples at several times the fundamental frequency.



```
(Time 0.)
  (Strings
    ((4 (G /2))
     (3 (D *1))
     (2 (A *1))
     (1 (E *2))))
  (Dynamic PP)
  (Tempo 0.9 <)
  (Stroke 4)
  (Place_and_Vibrato 4 FS *1)
(After 0.5)
  (Place_and_Vibrato 4 E *1)
  (Dynamic PP <)
(After 0.5)
  (Accent_and_Vibrato 4 D *1)
(After 1.0)
  (Martele 4)
  (Dynamic P)
(After 1.0)
  (Lift 4)
  (Place_and_Vibrato 3 FS *1)
  (Stroke 3)
  (Tempo 1.0 <)
(After 0.5)
  (Place_and_Vibrato 3 A *1)
(After 0.5)
  (Place_and_Vibrato 3 E *1)
  (Dynamic P >)
(After 2.)
  (Martele 3)
  (Place_and_Vibrato 3 FS *1)
  (Dynamic PP <)
  (Tempo 1.1 <)
(After 0.5)
  (Place 3 A *1)
(After 0.5)
  (Accent_and_Vibrato 2 B *1)
  (Stroke 2)
  (Open 3)
  (Light_Stroke 3)
  (Dynamic P)
(After 1.5)
  (Place 3 E *1)
  (Lift 2)
  (Dynamic PP)
(After 0.3)
  (Lift 3)
(After 0.2)
  (Place_and_Vibrato 4 E *1)
  (Martele 4)
  (Dynamic P >)
  (Tempo 1.2)
(After 0.75)
  (Place 4 D *1)
(After 0.25)
  (Accent_and_Vibrato 4 FS *1)
  (Dynamic PP >)
  (Vibrato_Excursion 0.02 >)
(After 1.)
  (Place_and_Vibrato 4 E *1)
(After 1.)
  (Lift 4)
  (Dynamic N)
  (Vibrato_Excursion 0.005)
(After 0.01)
(End)
```

Figure 2. Player system input notation.

A violin passage has been transcribed as input for the violin player system. The specific choices for fingering and bowing represent one of several possible interpretations for performing the score. Each event in the time-ordered list is specified by an articulation name, the string number and an (optional) new pitch to be sounded. Phrasing marks such as accents and dynamic changes are included.

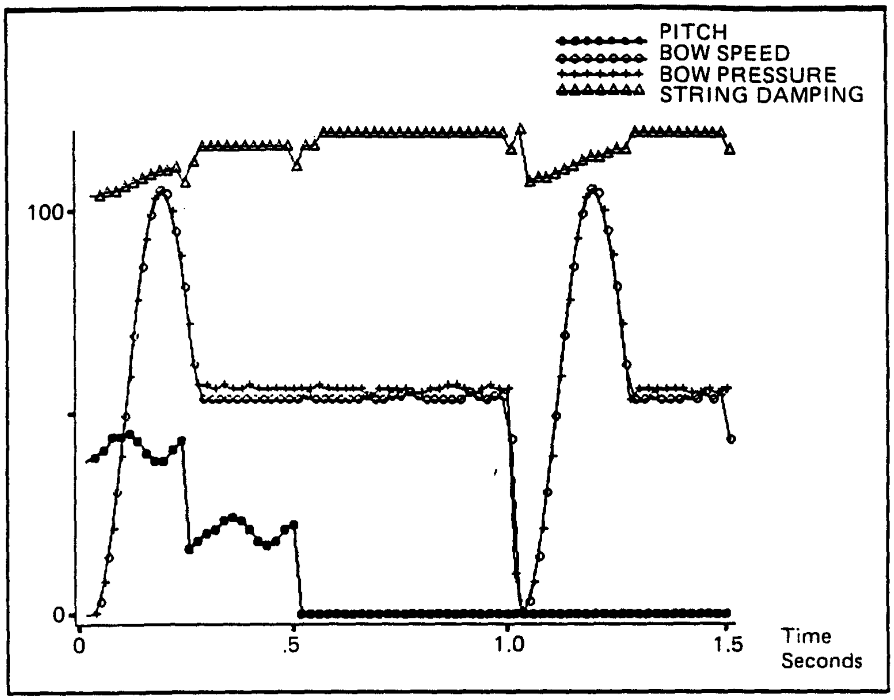


Figure 3
Envelopes generated by the player system. The plot provides a detailed view of envelopes for the synthesis of the first four notes of the previous score. To give an approximate idea of scale: the vibrato excursion shown in the graph is almost a semitone. Bow speed and pressure are shown in log units, from minimum to maximum effort divided into a scale of 127 values (as with MIDI control data). String damping is at a minimum at the very top of the scale, where it corresponds to a free ringing string. The graph begins with a three-note slur played in the left hand, followed by a repeated pitch in the bow. As the left-hand finger moves to a new pitch, damping increases slightly until the new note is fully sounded.

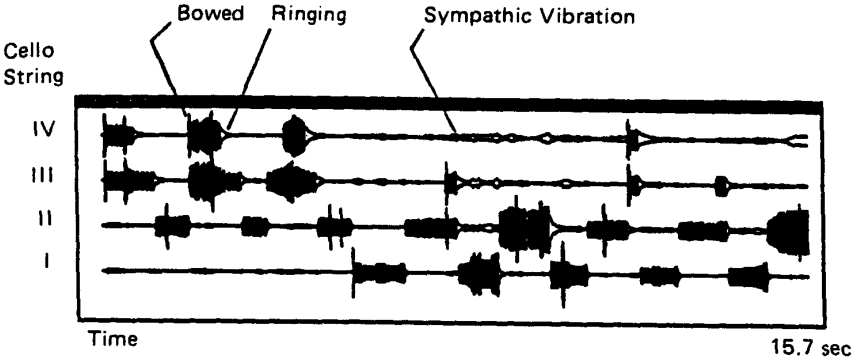
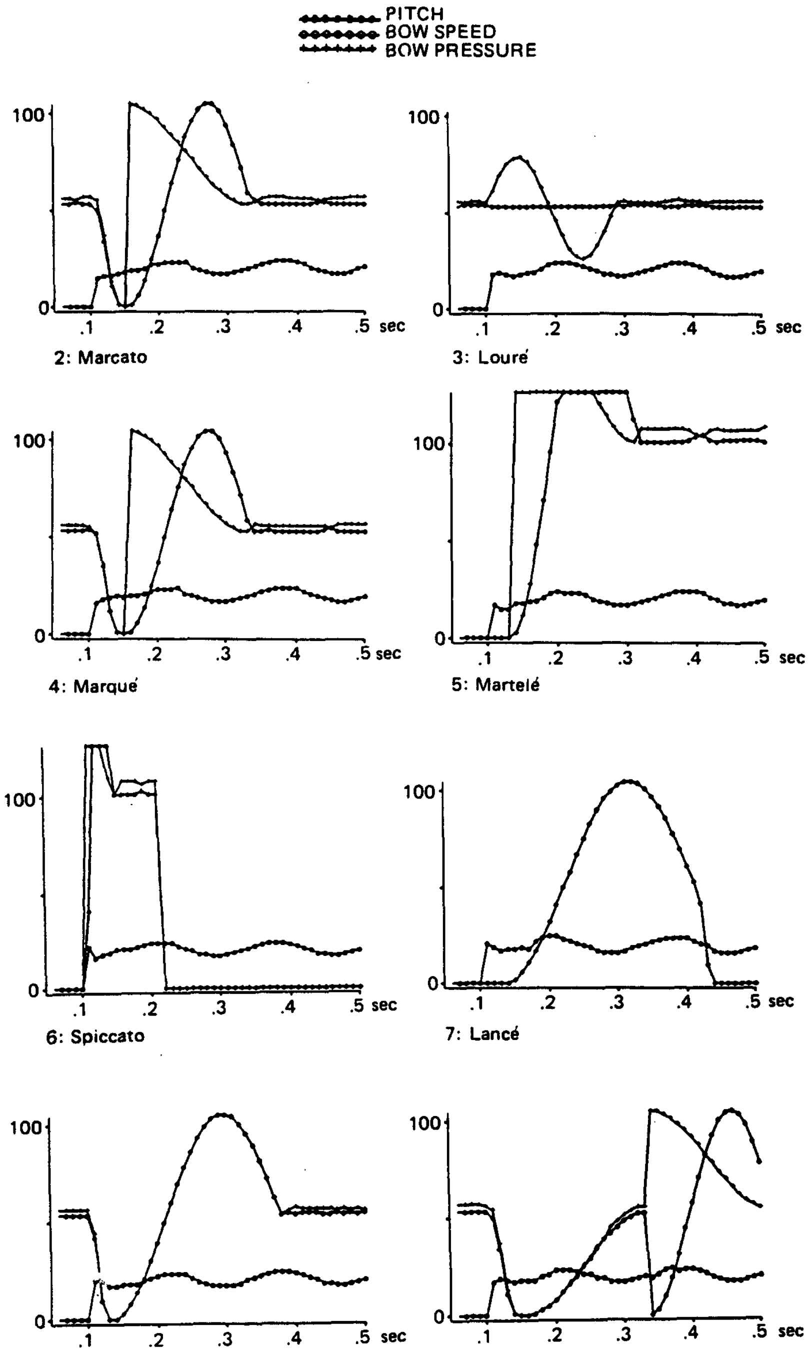


Figure 4
Synthesized string motion. Output level has been recorded for each string in isolation in the synthesis of sound example number 83, *Wrassled with a Wildcat.* Shading shows bowed portions, and open areas show either ringing or sympathetic vibration caused by bowing on other strings. The dynamic level of each double stop is coordinated by the global phrase controls.



2: Marcato

3: Louré

4: Marqué

5: Martelé

6: Spiccato

7: Lancé

8: Porté

9: "Impossible Porté"

Figure 5
Envelopes for various bowed articulations. These envelopes were generated in the process of synthesizing sound example number 85: the comparison of bowed articulations. Graphs show the second stroke of each group of four tones.