3rd Congress of the Alps Adria Acoustics Association
27–28 September 2007, Graz – Austria

LOCATION DETERMINATION OF ACOUSTICAL SOURCES USING A MICROPHONE ARRAY

Hannes Gamper, Thomas Walder, Alois Sontacchi

IEM Institute of Electronic Music and Acoustics, Graz, hannes.gamper@student.kug.ac.at

Abstract: *In case of mobile loudspeaker arrangements to enable virtual acoustic applications there is an essential need to know precisely the placement of each loudspeaker in the real room. Within common state-of-the-art 3D audio rendering approaches the real positions of the loudspeakers are directly related to the computed loudspeaker feeds. Beside deviations of ideal arrangements, permissible product tolerances, etc. digital audio signal processing can be used to compensate for.*

The proposed basic approach, based on the usage of a microphone array consisting of four elements, describes a convenient solution to determine the real position of any loudspeaker within the rendering arrangement in azimuth, elevation and distance in relation to a defined reference point. Caused by a centrally located microphone array, the complex compensation gains can be easily obtained, too.

Omitting a signal form a loudspeaker will cause different arrival times on the microphone array elements. Calculating time of arrival differences for each couple of elements will lead to a set of distinct delays.

To solve the problem to directly assign the obtained set of delays to the corresponding position of the regarded loudspeaker in the three dimensional space, additional information based on the synchronisation between the loudspeaker arrangement and the microphone array is applied.

Key words: acoustic location determination, microphone arrays, swept sine, parabolic interpolation

1. INTRODUCTION

Sound propagation generally speaking follows certain geometrical and physical rules, the inspection of which may shed light on the sound's source, i.e. information about its layout, its power and – its position.

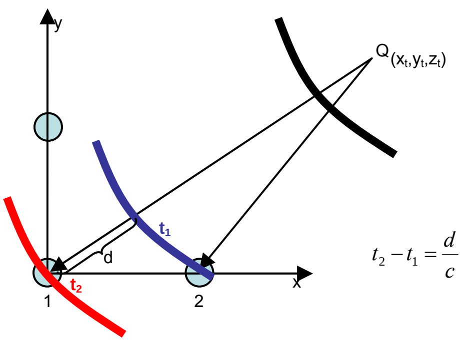

Fig.1. Time-Of-Arrival (TOA) differences caused by different transmission path lengths.

Nature provides us with an example of how localization of acoustical sources may be achieved: Our brain is able to determine the direction of a sound source making use of simple physics. The ear signals show interaural time and level differences, as a sound wave generally reaches both ears at a different time and with a different sound level. The human brain examines these differences and deduces the source's direction from them [1].

Pursuing an analogous technical approach, a sound source's direction – assuming a few simplifications – may be derived from the time difference of the signals arriving at two microphones at different positions (see figure 1).

In [2] a basic overview of various methods to estimate the position of a sound source is given. Beside the methods based on the time-of-arrival (TOA) differences even other properties like sound pressure differences, intensity differences are shown. However, in practical applications, the robustness towards other disturbing sound sources and the immunity against interfering reflections is of great importance. Therefore, methods based on the TOA approach exhibit more reliable performance.

As shown by S. S. Reddi in [3], to determine the exact position of a source in three dimensional space, it is necessary to examine the signals and the time differences respectively of at least four microphones. Based on Reddi's work, the method here presented will make use of an array of microphones to localize an acoustical source.

The proposed basic approach, based on the usage of a microphone array consisting of four elements, describes a convenient solution to determine the real position of any loudspeaker in azimuth, elevation and distance in relation to a defined reference point e.g. origin or sweet spot of the loudspeaker arrangement. Caused by the distinct centrally located arrangement of the microphone array elements, the complex compensation gains can be easily obtained, too.

Omitting a signal form a loudspeaker will cause different arrival times on the microphone array elements. Times of arrival can be computed by the usage of the cross-correlations of the direct path signals received at the microphone array elements and the transmitted loudspeaker signal. Calculating time of arrival differences for each couple of elements will lead to a set of distinct delays.

To solve the problem to directly assign the obtained set of delays to the corresponding position of the regarded loudspeaker in the three dimensional space, additional information based on the synchronisation between the loudspeaker arrangement and the microphone array is applied.

In an Higher Order Ambisonic setup (HOA, [4,5]), for example, the loudspeakers ideally are positioned on a hemisphere. If, however, their actual positions do not lie on a hemisphere, knowing their relative time delays can be used to "correct" their positions, i.e. virtually placing them on an acoustic hemisphere by applying an appropriate delay and a gain adjustment to each of the loudspeaker signals.

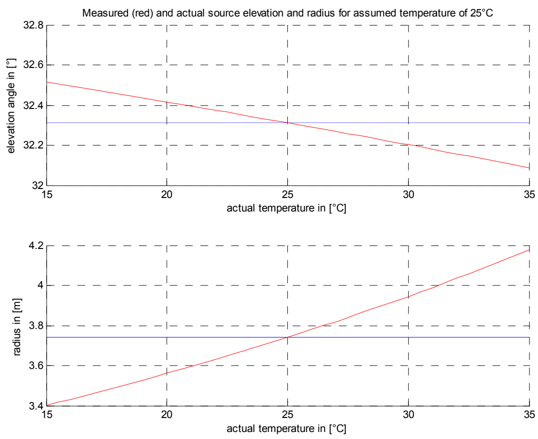

Fig.2. A temperature drift of ±10° affects the calculated source position (for sensor setup used see figure 4).

It has to be pointed out, that the correctness of the calculations relies upon the knowledge of the sound propagation speed. The latter being temperature dependant, it has to be calculated for the specific measurement environment:

$$c = 331 + 0.6 * T \qquad (1)$$

where c is the speed of sound and T is the room temperature.

Therefore a temperature shift may affect the calculations and lead to erroneous results (see figure 2)

In the case of a HOA system these deviations are negligible, as the calculated azimuth and elevation angles, which for such a system are of major importance, are only marginally affected.

When considering a system based on the absolute loudspeaker distances, such as a Wave Field Synthesis system (cf. [6]), however, it has to be assured that the speed of sound is precisely known, as it heavily affects the calculated distances.

To overcome the addressed problem in the latter case, a reference measurement shall be performed, e.g. by determining the acoustic traveling time on a known path length within the measurement area.

Furthermore, considering any of the mentioned multi-channel loudspeaker system, also the necessary gain adjustments are to be determined, in case the loudspeaker signals show different sound levels at the listening position.

2. MATHEMATICAL MODEL

To calculate the source's position, the method presented by Seenu S. Reddi in [3] is pursued.

Reddi proposes the use of one reference receiver and at least three other receivers (see figure 3), whose positions are to be linearly independent. Measuring the time differences of incoming signals between the receivers, the corresponding path length differences from the target to each of the receivers can be calculated. For a known geometrical layout of the receiver-array, this can be expressed in an equation system. Written in matrix form, it leads to a quadratic equation in R_t, the target distance:

$$R_t^2 \, (\boldsymbol{C}^T\boldsymbol{C} - 1) + R_t \, (\boldsymbol{C}^T\boldsymbol{Y} + \boldsymbol{Y}^T\boldsymbol{C}) + \boldsymbol{Y}^T\boldsymbol{Y} = 0 \qquad (2)$$

Solving this equation for R_t, with the use of the help vectors $\boldsymbol{C}$ and $\boldsymbol{Y}$, two solutions are found for V, the vector containing the target coordinates:

$$\boldsymbol{V} = R_t\boldsymbol{C} + \boldsymbol{Y} \qquad (3)$$

one of which corresponds to the actual position of the source. The help vectors $\boldsymbol{C}$ and $\boldsymbol{Y}$ depend on the actual sensor positions and on the calculated TOA differences. A detailed description is given in [3].

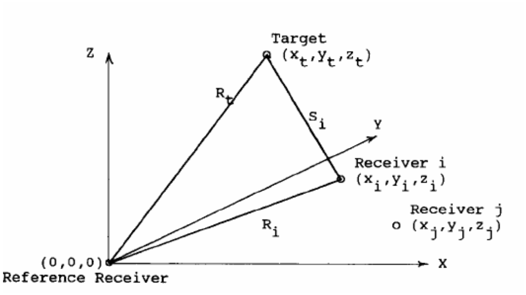

Fig.3. Reddi's method & parameter description (cf. [3])

3. TEST SETUP

A possible scenario to localize an acoustical source is depicted in figure 4. The position of a loudspeaker is to be determined, using an array of (at least) four microphones. Loudspeaker and microphones are connected to and controlled by a computer. An application provides the test signal for the loudspeaker and records the microphone input.

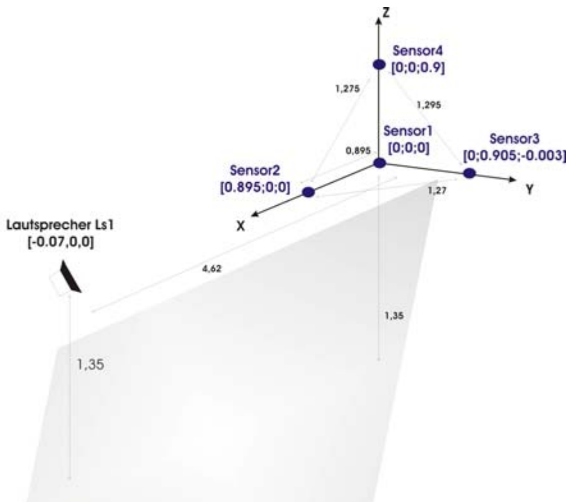

Fig.4. Sketch of a possible test scenario.

For the array to be easy to set up and in order to simplify the latter calculations, in the experimental setup only the minimum of four microphones is used. The reference microphone is put at the origin of a three dimensional Cartesian coordinate system, the other three microphones each on one of the x-, y- and z-axis in equal distances (e.g. $L=1m$) to the origin, thus ensuring linear independence of their positions. Ideally a mounted array configuration is used. In [2] the appropriate choice of distance L is handled.

Usually the array should be placed into the listening area or "sweet spot" of the loudspeaker system, as all results are presented relative to the array's position. Also in an ideal setup all loudspeakers are directed to this area, hence influence of room reflections and reverberation are minimized.

3.1. Time delay determination

The sound waves emitted from the loudspeaker reach the microphones at different times, due to the differing path lengths from the source to each of the microphones. As for the loudspeaker signal, any de-correlated signal may be used – in the described test setup, it consisted of a logarithmic sweep, as presented by A. Farina in [7], which provides good signal to noise ratio and immunity against harmonic distortion caused for instance by the loudspeaker as well as early reflections and room reverberation.

By convolving the microphone signals with the time-reversal of the excitation signal, the impulse response is obtained for each microphone.

To eliminate temporal smearing of the signal caused by the loudspeaker's temporal response behavior, a high-pass filter may be applied to the obtained impulse responses.

Furthermore, to eliminate interfering early reflections and to reduce the effect of the room reverberation only the direct part of the impulse response is processed. Therefore, a half Hanning curve is used to fade out the impulse responses approximately one millisecond after their maximum. From these direct parts, the cross-correlations are calculated to determine the relative delays between the sensors. The offset of the cross-correlation maximum corresponds to the time difference between two microphone signals, or the path length difference respectively (see figure 5).

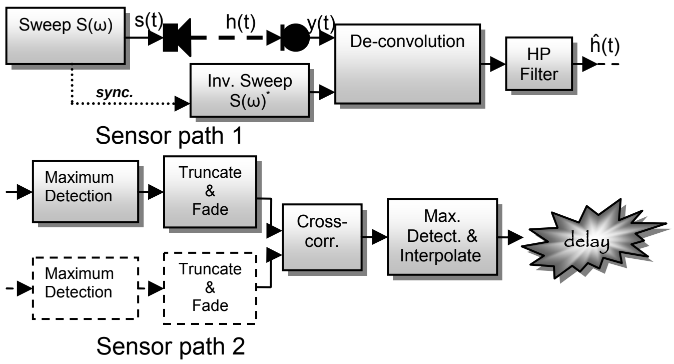

Fig.5. Block diagram of delay determination using swept sine impulse response measurement and crosscorrelation.

Caused by the fact, that the sampling rate fs determines the quantization factor of the measured delay, it should be chosen as high as possible (e.g. fs@44.1kHz & c=340m/s provides a spatial resolution of 0.7cm). :

$$\Delta t=1/fs \tag{4}$$

As a precise determination of the delays is vital for acceptable results, an interpolation is built into the application to increase the resolution to up to a few hundredths of one sample.

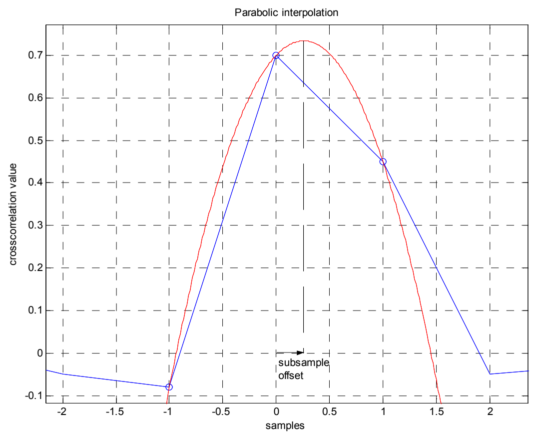

Fig.6. Parabolic interpolation

A straightforward method to achieve this is the parabolic interpolation (cf. [8]), where a parabola is placed on three points of the cross-correlation function, i.e. its maximum and the two neighboring values. The parabola is herewith uniquely described. Its maximum provides a sub-sample estimate of the real cross-correlation maximum (see figure 8).

3.2. Gain adjustment

To ensure equal gains of all loudspeakers in the system, the mean sound pressure level produced in the listening area by each loudspeaker, which will vary according to the loudspeaker's position and/or gain setting, has to be measured.
From the energetic summation of the impulse responses gathered from all sensors, a local and temporal average of the sound pressure level in the listening area can be deduced for each loudspeaker. It indicates the necessary gain adjustment to attain an optimal setup consisting of (virtually) equidistant sound sources with equal gains.

4. SOLUTION(S)

With the knowledge of the delays between the sensors in the array, the equation system is set up and solved as described in [2], to attain the source's coordinates.
As mentioned before, the method used to solve the equation system provides two solutions. Up to now no appropriate algorithm could be found to uniquely identify the valid one of the two solutions. It is therefore up to the person performing the measurements to decide which result to choose (see figure **7**-9).
To allow for a quick visual estimation of the relevance and correctness of the solution, the calculated position is presented as distance, azimuth, and elevation angle relative to the reference microphone, or the centre of the listening area respectively.

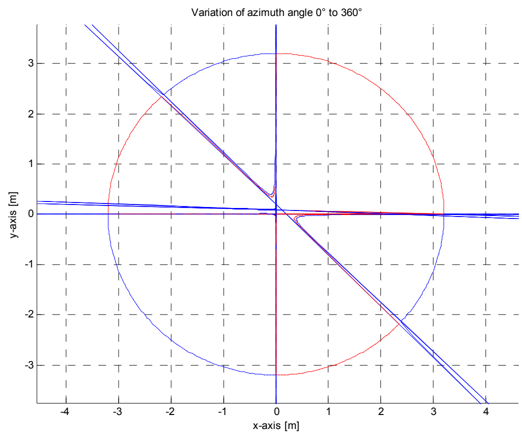

Fig.7. Azimuth locations of the two obtained solutions (blue: 1st solution, red: 2nd solution)

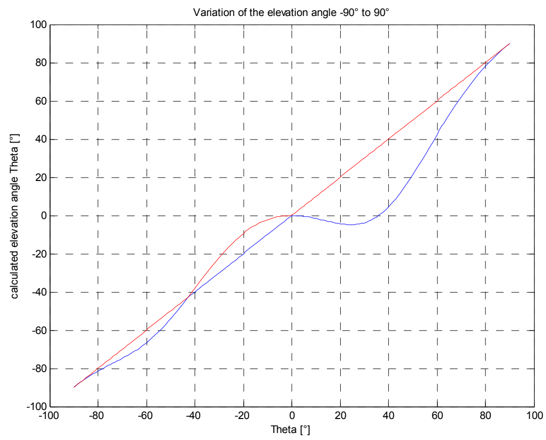

Fig.8. Elevation of the two obtained solutions (blue: 1st solution, red: 2nd solution)

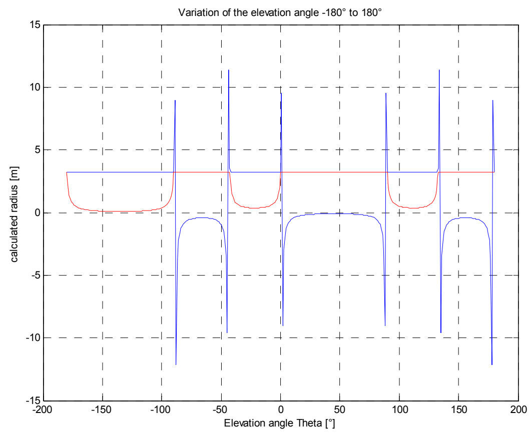

Fig.9. Radii of the two obtained solutions (blue: 1st solution, red: 2nd solution)

Caused by the fact, that our sound source is synchronized to the receivers, we can overcome this drawback.
Therefore, we measure the absolute acoustic delay between the loudspeaker and the reference microphone, and by comparing it with the calculated source distance of both solutions we can select the appropriate one. The appropriate solution is determined by the minimal deviation of the absolute difference of the measured and solved source distance (radii).

5. RESTRICTIONS AND LIMITATIONS

The most obvious drawback of the presented method is the fact, that for a fixed source position the choice of the right solution generally can only be done via visual estimation – i.e. by discarding the less "probable" solution for the source position.
One possibility to get around this problem might be to vary the array's position and observe the change of the calculated source position: Only one solution will smoothly change corresponding to the array's movement – and thus be selected as the "right" solution by an algorithm (see figure 7 - 9).
In the transition area, however, the valid solution switches from one to the other. Here both solutions show very close results, and to exclude one via visual estimation or an algorithm might be very difficult or even impossible. On the one hand this implies that both solutions provide a more or less acceptable result, on the other hand though this fact lowers the quality of the measurement, in case the solution further off the actual source position is picked.
Another difficulty encountered is the fact, that the quality of the results in general – and especially the calculation of the distance – is highly dependant upon the exactness of the measured time differences between the microphones. Measurement errors in the range of one sample (caused for example by a deviance of one microphone from the calculated position) can noticably affect the calculation results.

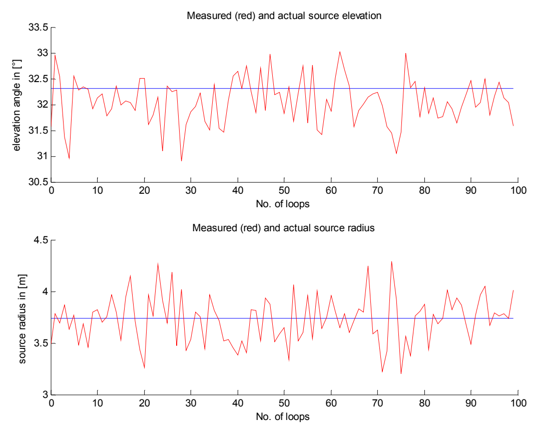

Fig.10. Position deviation caused by little time jitter (time jitter < 1sample, fs@44.1kHz)

In a simulation**,** normally distributed white noise with mean zero, variance one and standard deviation one sample was added to the calculated delays for each of a total of 100 calculations of the source position for the same fixed source, the outcome being that especially the results for the source radius are remarkably affected by these tiny time delay deviations, whereas the calculated elevation differs only slightly from the actual value (see figure 10).
In order to provide an acceptable solution for the source's distance we need additional information concerning the true distance of the measured loudspeaker position.
To obtain the absolute "acoustic delay" between the reference sensor and the investigated loudspeaker, we calculate, based on the modified impulse responses (cf. chapter 3), the Energy Decay Curve (EDC).
We define the delay to be measured at that time point, where the EDC falls below 90%, or by approximately 0.5 dB respectively. Alternatively, the delay can also be set to be the energy balance centre of the modified impulse responses.

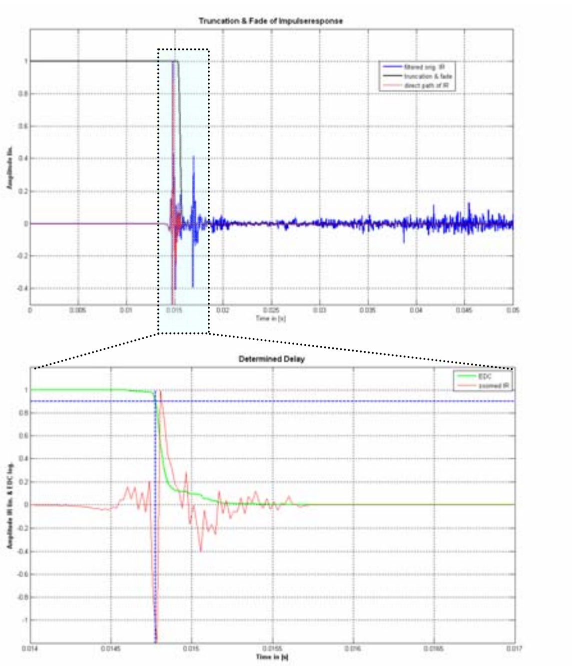

Fig.11. Delay determination calculating the impulse response's Energy Decay Curve (EDC). The graph above shows the direct part of the impulse response and the eliminated reflections.

After determining and subtracting the system inherent delay caused by the computer's audio buffers and processing time, the sound's path length can be derived, providing an excellent estimation for the distance of the loudspeaker's acoustical centre from the reference microphone.
As already mentioned, the result of this measurement for the distance may also provide a possible criteria to exclude on of the two solutions for the source position: the one solution nearer to the calculated source distance

may be accepted as the valid solution. Furthermore, based on the measured acoustic delays the relative deviations of the loudspeakers can be compensated.

6. CONCLUSION

The precision of the measurements plays a major role when trying to exactly localize a target by calculating its distance to the known positions of some receivers. Minimal errors already may affect the results in a way to make them useless. Thus the practical usability of the presented method to automatically determine the position of an acoustical source highly depends upon the exact positioning (or knowledge of position) of the microphone array elements. Therefore, a fixed microphone arrangement should be used.

On the other hand, the location determination can be done in "real-time", with a relatively small effort in terms of installation and preparation, especially when repeating the measurement for different sources within the same system, making it a useful tool when multiple loudspeaker positions have to be determined quickly. This may be a huge advantage particularly when the loudspeakers are mounted in places difficult to reach, e.g. under a balcony etc, or when the setup is mobile and thus subject to frequent changes in terms of positioning and layout.

It should be pointed out, however, that the method provides still a great potential for developments and improvements. From the mathematical model any array geometry is allowed, with any number of microphones greater than or equal to four. Thus different array layouts could be defined, optimizing the resolution of the array, and minimizing measurement errors.

In order to automatically determine all relevant positions in a multi-channel loudspeaker system, the application has to be able to automatically exclude one of the two solutions when determining the location of a target. Possibly array geometries other than the presented one could simplify this problem.

If the application is combined with the loudspeaker control system, more specific features may be developed, for example an automatic delay and gain adjustment, able to simulate an acoustical hemisphere, i.e. by positioning the loudspeakers of the system on a virtual hemisphere, or any arbitrary layout.

Further measurements could take into account also the loudspeakers' impulse responses and set equalizers to "correct" or balance them.

REFERENCES

[1] J. Blauert: ***Spatial Hearing*** - 2nd ed., MIT Press, Cambridge, MA, 1997

[2] W. Würfel: **Passive akustische Lokalisation,** Diploma Thesis at the IEM, TU Graz, Austria, 1997

[3] S. S. Reddi: **An Exact Solution to Range Computation with Time Delay Information for Arbitrary Array Geometries**, *IEEE Transactions on Signal Processing, Vol. 41, No. 1*, 485-486, January 1993

[4] M. A. Gerzon, **Periphony: With-Height Sound Reproduction**, *J. Audio Eng. Soc.*, Vol. 21(1), pp. 2-10, 1973.

[5] Zmölnig J., Ritsch W., Sontacchi A. **The IEM CUBE**, *ICAD - International Conference on Auditory Display*, Boston University, Boston, MA, July 7-9, 2003

[6] E. Verheijen, **Sound Reproduction by Wave Field Synthesis**, PhD Thesis, TU Delft, Netherlands, 1998.

[7] A. Farina: **Simultaneous Measurement of Impulse Response and Distortion with a Swept-Sine Technique**, *Presented at the 108th AES convention*, Paris 18-22 February, 2000

[8] X. Lai, H. Torp: **Interpolation Methods for Time-Delay Estimation Using Cross-Correlation Method for Blood Velocity Measurement**, *JIEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, Vol. 46, No. 2,* 277-290, March 1999